INTERNATIONAL COMPARISONS OF REAL INCOME

By

Robert James Ackland

A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
THE AUSTRALIAN NATIONAL UNIVERSITY
14TH MARCH, 2001

THE AUSTRALIAN NATIONAL UNIVERSITY

SCHOOL OF ECONOMICS
FACULTY OF ECONOMICS AND COMMERCE

I the undersigned hereby certify that the following thesis entitled "**International Comparisons of Real Income**", submitted for the degree of **Doctor of Philosophy** at **The Australian National University** is my own original work.

Signature of author: R J Ackland
Robert James Ackland

Dated: 14th March, 2001

THE AUSTRALIAN NATIONAL UNIVERSITY

Date: **14th March, 2001**

Author:	**Robert James Ackland**
Title:	**International Comparisons of Real Income**
School:	**Economics**
Faculty:	**Economics and Commerce**
Degree:	**Ph.D.**
Supervisor:	**Steve Dowrick**

Permission is herewith granted to The Australian National University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

R J Ackland

Signature of Author

To Mum - your love and encouragement has been constant, even when facing your own difficult challenges.

Table of Contents

List of Tables

List of Figures

Preface

I wrote this thesis while at the A.N.U. between 1997 and 2001, but I really began the "PhD process" back in 1993 when I commenced graduate study in the U.S. My studies in the U.S. and also the time I spent at the World Bank in Washington, D.C. furthered my training in economics and helped me to decide on the areas of research I was interested in. If, as an undergraduate, I had known that I would write a PhD thesis on index number theory, I wouldn't have believed it. When I arrived at the A.N.U. in 1997, my supervisor Steve Dowrick said he would be delighted if I worked on international comparisons, which was one of the areas he was researching on. On first reading some articles in the area, I didn't think it was for me, as at the time I was more interested in applied welfare analysis using microdata. However, international comparisons (and index number theory in general) is an area that grows on you, and it offers the researcher a great mix of microeconomic theory and empirical applications. I have found the experience of writing this thesis very rewarding, and I hope the reader is similarly rewarded.

Acknowledgements

First, I would like to thank my family for all their support during my PhD studies. I would also like to thank Kazuko who has been there for me during the latter stages.

I would like to thank Steve Dowrick for always being available for discussion of research ideas and for offering helpful support and advice, both academic and non-academic. Steve has reinforced for me a belief that research can and should be a fun journey of discovery (even on as dry a topic as index number theory).

I would like to thank Paul Chen for carefully reading several versions of different chapters and offering lots of useful advice. Jeff Kline and John Quiggin also read a first draft and provided helpful comments. Jeff Borland provided useful comments on an early version of Chapter 6, and he has also been a constant source of good advice and help ever since my Honours year at Melbourne University.

I would like to thank Peter Neary and Ian Crawford for inviting me to present papers at University College Dublin and the Institute for Fiscal Studies, respectively. These visits and on-going contact after the visits were very helpful for me. I would also like to thank Hal. R. Varian for directing me to his NONPAR program, which was very useful for my work in Chapter 2.

There are many other people who have helped me along the PhD process, either academically or through work or professional experience. I hope this list is not too incomplete: Jeanine Braithwaite, Bob Breunig, Trevor Breusch, Jane Falkingham, Simon Grant, Jesko Hentschel, Jenny Hunt, Peter Lanjouw, Jenny Lanjouw, Phillip Leslie, Branko Milanovic, Mike Mills, David Prentice, T. Paul Schultz, Kinnon Scott, Rod Tyers, Lynne Williams.

I would like to thank participants in seminars at: The A.N.U.; Institute for Fiscal Studies; University College Dublin; 1999 PhD Conference in Economics and Business (U.W.A); The Melbourne Institute of Applied Economic and Social Research.

The financial support of the Australian Postgraduate Award and the Fulbright Association are appreciated, as is the support of the School of Economics, Faculties of Economics and Commerce, at the A.N.U.

Software used for empirical work and thesis production

The following software packages were used for the empirical work: GNU C (in Linux), Excel, STATA and Gauss. The thesis was typeset with LaTeXusing the MiKTeX distribution for Windows and WinEdt as the shell editor. Indexing and bibliographic management were performed using the `makeindex` and `bibtex` tools, respectively. Graphics were produced using Xfig and Gnuplot (in the Linux environment). I used Xemacs as the shell editor for the C programming. Several of the above software tools have been provided to users free of charge - I would like say a special thanks to authors of these packages.

Canberra
March, 2001

Abstract

In this thesis, an analysis of the use of multilateral true welfare indexes in cross-country welfare comparisons is presented. There are two main goals of the study. The first aim is to present a comprehensive review of different approaches to conducting cross-country welfare comparisons, with particular emphasis on the Revealed Preference (RP) approach to welfare measurement. In Chapter 1, a review of the different approaches to index number construction is presented. The RP approach to welfare measurement is reviewed, and the construction of bilateral and multilateral true welfare indexes under this approach is introduced. In Chapter 2, RP tests for common general preferences are reviewed and three approaches to constructing bilateral true welfare indexes are outlined. In Chapter 3, results on the existence and construction of multilateral true welfare indexes are presented and the Ideal Afriat Index of Dowrick and Quiggin (1997) is constructed using 1980 and 1993 data from the International Comparison Programme (ICP).

The second aim of the thesis is to present three extensions on the use of multilateral true indexes in cross-country welfare comparisons, and to apply these extensions to the ICP data. One of the criticism of the standard RP approach to welfare measurement is that it is non-stochastic in that it does not allow for the existence of measurement or consumer optimisation error. In Chapter 4, the RP approach is extended to enable approximate multilateral welfare comparisons. It is found that under moderate levels of consumer optimisation error, the pairwise rankings between certain countries change, thus suggesting that these countries should perhaps be ranked as equivalent.

Homotheticity is necessary and sufficient for the existence of a unique (i.e. base-country invariant) welfare index. However, homotheticity is very restrictive and in Chapter 5, results on the construction of unique marginal welfare indexes, which exist under affine homothetic preferences (of which, homotheticity is a special case) and measure the utility gained from supernumerary consumption, are presented. A particular multilateral true marginal index, the Ideal Afriat Marginal Index, is proposed and it is shown that the Ideal Afriat Index is a special case of the Ideal Afriat Marginal Index.

The Ideal Afriat Marginal Index is constructed for a particular subsistence consumption bundle. For a given data set, an iterative procedure can be used to find the range within which the subsistence bundles must lie (for a homothetic data set, one of the subsistence bundles will be a vector of zeros). By taking a sample of possible subsistence bundles we can construct bounds to the Ideal Afriat Marginal Indexes and bounds to the rankings implied by these indexes. For the ICP data, it is found that there are inconsistencies in the country rankings implied by the Ideal Afriat Index in that there exists another (equally valid) minimum subsistence bundle that results in different rankings. It is also shown that a comparison of the Ideal Afriat Index and the bounds to the Ideal Afriat Marginal Indexes can be used to identify poor and non-poor countries.

The majority of research on cross-country comparisons of welfare are based on per capita expenditure on goods and services; such measures of economic welfare ignore an important commodity, leisure, which a consumer purchases implicitly by not working. In Chapter 6 the joint commodity demand-labour supply framework is used to extend the RP approach to include the value of leisure time. A leisure-inclusive multilateral true welfare index is defined, and it is found that for the 1993 ICP data, the inclusion of leisure can have a marked impact on the welfare rankings of countries.

Introduction

> "The ground has been trodden so often that what one makes of it might be a personal accident." Sidney N. Afriat (1987, p.195) on index number theory.

There is a considerable body of research on the methods for constructing measures of real income for use in cross-country living standards comparisons. This research has partly been driven by the availability of cross-country data sets which contain comparable and highly disaggregated price and quantity data for countries from the entire development spectrum (the International Comparison Programme, ICP, is a good example of such a data source). Another factor contributing to continuing research on aggregation methods in the context of cross-country living standards comparisons is the demand for the comparative data that these methods produce.

For example, rankings of countries in terms of GDP per capita are used as an indicator of comparative economic performance, and thus evidence of the success or failure of different policy regimes. The measurement of convergence (or lack thereof) in countries' productivity and living standards is central to distinguishing between the neoclassical growth model and models of endogenous growth; such measurement requires as input consistent cross-country real income data. Cross-country rankings of living standards are also of direct interest to multilateral lending organisations, and are an important input into the determination of aid flows to developing countries.

However, one can argue that the major reason for the enduring popularity of research into cross-country welfare comparisons is that it is an application of the intrinsically fascinating area of index number theory. Index number theory is one of the oldest fields in economics, and has attracted the attention of some of the brightest minds in the discipline (and the rest of us). The body of research on index number theory is so large, that the above quote of Sidney N. Afriat seems particularly pertinent to one who is trying to make an original contribution to the field.

There are two main aims of this thesis. The first is to present a comprehensive review of different approaches to conducting cross-country welfare comparisons. This review, which is contained in the first three chapters, covers the major results in index number theory which are relevant to cross-country welfare comparisons. Further, there is an attempt to present some of the more mathematically challenging results in index number theory in a form which is accessible to a wider audience, and thus hopefully reduce the "start-up costs" associated with beginning research in this area. This review has been particularly rewarding to the author, who himself found some difficulty in understanding some of the source material. Indeed one could add to Afriat's quote: "The path that Afriat trod is difficult to follow, and success in this might be a personal accident."

While the material in Chapters 1-3 is essentially a review of previous work in the field, there are several innovative aspects of this review. In particular, a new typology which identifies the Revealed Preference (RP) approach to welfare measurement as an example of the economic approach to index number construction is presented. Under the RP approach, the existence of a representative consumer is tested for using revealed preference relations and true welfare indexes are constructed when the hypothesis that data are rationalised by a utility function cannot be rejected. The RP approach to index number construction has mainly been used in the context of within-country cost-of-living measurement. However, the RP approach is increasingly being used in cross-country welfare comparisons - the Ideal Afriat Index of Dowrick and Quiggin (1997) is an example.

There has been a lot of research on the RP approach; one of the aims of the review chapters is to incorporate the main results into a consistent framework which will hopefully be of use to other researchers working in this area. The assumptions required to construct within-country cost-of-living indexes using the RP approach are different to those necessary for using this approach to compare welfare across countries; in this review the assumptions and methods used in each context are clearly stated. Further, an attempt has been made to precisely define the different terms which are encountered in the RP approach. For example, precise definitions are given to the terms: welfare index, true welfare index, bilateral index, multilateral index, bilateral true index, multilateral true index.

The second major aim of the thesis is to present several theoretical extensions to the RP approach to cross-country welfare comparisons, and to apply these results to 1980 and 1993 data from the ICP. The starting point for these extensions is the Ideal Afriat Index which Dowrick and Quiggin (1997) proposed as an appropriate index for multilateral welfare comparisons. In Chapter 4, the RP approach is extended to allow approximate multilateral welfare comparisons and an Ideal Afriat Index which incorporates specific levels of consumer optimisation error is constructed.

The Ideal Afriat Index is a true index which exists when preferences are homothetic. However, homotheticity is a very restrictive assumption, and may not be appropriate in a cross-country context (where incomes typically vary markedly). In Chapter 5, the Ideal Afriat Marginal Index is proposed and constructed using the ICP data. This index is constructed under the less restrictive assumption of affine homotheticity, and it is shown that the Ideal Afriat Index is in fact a special case of the Ideal Afriat Marginal Index. It is further shown that the Ideal Afriat Index and the bounds to the Ideal Afriat Marginal Index can be used to classify countries as poor or non-poor, thus avoiding arbitrary decisions about where international poverty lines should be set.

In Chapter 6, the joint commodity demand-labour supply framework is used to extend the RP approach to include the value of leisure time in cross-country welfare comparisons. The Ideal Afriat Index is adapted for use in leisure-inclusive cross-country welfare comparisons, and it is found that the inclusion of leisure has a marked impact on the welfare rankings of countries.

A more detailed description of the chapter contents follows.

Chapter 1

One of the main objectives of Chapter 1 is to define the multilateral true welfare index (of which, the Ideal Afriat Index is an example) which is the focus of the rest of the thesis. A multilateral true index is constructed under the RP approach to welfare measurement, which itself is an example of the economic approach to index number construction (and is thus based on the utility maximisation framework).

The definition of a multilateral true index used in this thesis is that it is an index which has two properties. First, it is a multilateral index, and second, it is a true index. A multilateral index is defined as an index which satisfies the property of circularity: the real income of country i relative to country j is the same whether the two are compared directly or via an arbitrary intermediate third country k. It can be shown that circularity implies base-country invariance, but the reverse is not true. Examples of multilateral indexes are the superlative indexes (the EKS and CCD indexes) and also the Geary index. A bilateral index is defined as any index which is not a multilateral index and examples are the Laspeyres and Paasche quantity indexes and the Ideal Fisher Index.

The definition of a true index is more involved. It can be shown that if the demand data can be rationalised by a representative consumer, then there exist money-metric welfare measures called the Allen welfare indexes which are constructed using the (unobservable) expenditure function. With general preferences, a welfare comparison between two countries can be made using either country's price vector. Hence base-country invariance does not hold and there is an issue of *non-uniqueness* - the welfare comparison can be made using either the base-weighted or current-weighted Allen welfare index. Additionally, since we do not have complete knowledge of preferences, the Allen welfare indexes are unobservable, that is, there is an issue of *indeterminateness.*

One of the innovative aspects of Chapter 1 is that it is shown how we can construct observable bounds to the Allen welfare indexes, which are called the classical and fixed-weight bounds. Since the unobservable Allen welfare indexes must be contained within these bounds, a true welfare index is defined as a set of numbers that satisfies these bounds (contingent on there existing a utility function that rationalises the data). In Chapter 2, RP tests for common preferences are presented and three methods for constructing bilateral true welfare indexes are reviewed.

If we cannot reject the hypothesis that the demand data are rationalised by a homothetic utility function, then the Allen welfare indexes coincide and there exists a unique welfare index. Thus, with homotheticity, there is no longer a problem of non-uniqueness, however the problem of indeterminateness remains. However, it can be shown that the unique welfare index must be contained within bounds provided by the Paasche and Laspeyres quantity indexes. This leads to the definition of the unique true welfare index, namely it is a set of numbers the ratios of which are contained within the Paasche-Laspeyres (P-L) bounds. It follows that a test for the consistency of a given data set with homotheticity is equivalent to the test of the existence of a unique true welfare index.

It is shown in Chapter 1 that a unique true welfare index will satisfy circularity, and hence is a multilateral true index. While the superlative and Geary indexes are multilateral indexes, they are not multilateral true indexes since it is possible that they will lie outside the P-L bounds. Tests for the existence of multilateral true welfare indexes and their construction are reviewed in Chapter 3.

Chapter 2

In Chapter 2, a comprehensive review of the construction of bilateral true welfare indexes is provided. First, there is a review of the tests for common general preferences which originated with the work of Afriat (1967), Diewert (1973) and Varian (1982). The most easily implemented of these tests is the Generalised Axiom of Revealed Preference (GARP). For a given data set that satisfies the test of common preferences, the Allen welfare indexes exist, however the classical and fixed-weight bounds to these indexes tend to not be tight. In this chapter, three RP methods for constructing improved bounds to the bilateral true indexes are reviewed. The three methods for constructing bilateral true welfare indexes are: the GARP bounds of Varian (1982), the improved GARP bounds (which use expansion path information) of Blundell, Browning, and Crawford (1998) and the Afriat envelope bounds of Chavas and Cox (1997). One of the innovations of this thesis is the development of a framework which can be used to assess the extent to which the bounds to the Allen welfare indexes are tightened by each of these methods.

Chapter 3

In Chapter 3, the tests for homothetic common preferences (Afriat (1972, 1981), (Diewert 1973) and (Varian 1983)), the most easily implemented of which is the Homothetic Axiom of Revealed Preference (HARP), are reviewed. The construction of multilateral true indexes is then summarised, and a distinction is made between multilateral true indexes which are used for bilateral comparisons (e.g. temporal studies where prices are compared to a particular base year) and multilateral true indexes which are used for multilateral comparisons (e.g. comparing the welfare of a particular country with that of the sample mean). An example of the latter is the Ideal Afriat Index of Dowrick and Quiggin (1997), and this index (and variants thereof) is the focus of the empirical work in the remainder of the thesis. The Ideal Afriat Index is constructed using 1980 and 1993 ICP data.

Chapter 4

One of the criticisms of the standard RP approach to welfare measurement is that it is an "all or nothing" approach in that it does not allow for the existence of measurement or consumer optimisation error, which may lead to the failure of the test of common preferences. In Chapter 4, the RP approach is extended to enable approximate multilateral welfare comparisons. In the first part of this chapter, the Afriat envelope approach reviewed in Chapter 2 is used to impute utility for countries which are found to not satisfy homothetic preferences. While this method allows for utility consistent multilateral comparisons across countries which do not share common homothetic preferences, it does not address the source of the failure of homotheticity. In the second part of Chapter 4, work of Afriat(1972, 1987) and Varian (1993) is extended to the construction of Ideal Afriat Indexes which incorporate specific levels of consumer optimisation error.

For the 1980 ICP data, it is shown that while 42 countries satisfy HARP, if one allows for 4.2 percent consumer optimisation error, then the welfare of 59 countries can be compared using an approximate version of the Ideal Afriat Index. This approach also allows for the identification of countries for which their pairwise ranking changes with different levels of consumer optimisation error. For the 1980 data, two such pairs of countries are identified (Denmark and Germany, and Belgium and Luxembourg), and it is argued that these countries should perhaps be ranked as equivalent. For the 1993 ICP data, 19 of the 24 countries satisfy HARP but with 2 percent consumer optimisation error, all countries are found to share common (approximate) homothetic preferences. Further, it is found that the rankings of Australia, Italy and the U.K. change when optimisation error is allowed for, thus suggesting that the rankings of these countries are perhaps indeterminate.

Chapter 5

Homotheticity is necessary and sufficient for the existence of a multilateral true welfare index. However, homotheticity implies restrictions on consumer behaviour which many find untenable, especially in a cross-country context. In Chapter 5, the second main extension to the RP approach to cross country welfare comparisons is presented. In this chapter, early work by Afriat (1972, 1977, 1987) on the construction of bilateral marginal welfare indexes (which exist when preferences are quasi homothetic, that is, the expansion paths are linear) is first reviewed. It is shown that, analogously with the existence of the unique welfare index, a unique marginal welfare index (which measure the utility gained from consumption in excess of a predetermined minimum subsistence bundle) exists when preferences are affine homothetic.

Affine homotheticity is a special case of quasi homotheticity where the linear expansion paths originate from a particular point, which can be identified as the minimum subsistence consumption bundle. Thus it can be seen that since homotheticity is a special case of affine homotheticity (where the minimum subsistence bundle is a vector of zeros, the origin) the unique welfare index is in fact a special case of the unique marginal welfare index. While affine homotheticity is still very restrictive, it is shown that this is the only relaxation from homotheticity which is consistent with the existence of a unique (marginal) welfare index.

Given knowledge of the minimum subsistence bundle, it is therefore possible to construct multilateral true marginal welfare indexes, which can be used for more general comparisons of welfare. A particular index, the Ideal Afriat Marginal Index is proposed for cross-country comparisons of marginal welfare. However, an obvious question is: how do we determine the minimum subsistence bundle? It is shown that for any given data set satisfying HARP there will exist a set of minimum subsistence bundles $\mathcal{G}$ and one of the vectors in this set will be the homothetic subsistence bundle (a vector of zeros). There is no empirical reason to base welfare comparisons on any particular bundle in $\mathcal{G}$, and hence the *bounds* to the Ideal Afriat Marginal Index are constructed using a sample of the subsistence bundles in $\mathcal{G}$. For a given data set satisfying HARP, a comparison of the Ideal Afriat Index and the bounds to Ideal Afriat Marginal Indexes provides a method for classifying countries as rich and poor. In particular, a country i can be classified as rich (poor) if the element of the Ideal Afriat Index corresponding to that country is equal to the relevant lower (upper) bound of the Ideal Afriat Marginal Indexes.

These methods are applied to the ICP data and it is found that for the 1980 data, $\mathcal{G}$ does not contain vectors of significant size and hence homotheticity is a very reasonable assumption. Despite this, a comparison of the (approximate) Ideal Afriat Index with the bounds to the (approximate) Ideal Afriat Marginal Index reveals that the pairwise ranking of 5 pairs of countries is dependent on the which subsistence bundle in $\mathcal{G}$ is chosen for the comparison, and hence these countries should perhaps be ranked as equivalent. It is also found that of the 59 countries for which it is possible to make (approximate) true welfare comparisons, 17 (or 28.8 percent) are classified as poor.

For the 1993 ICP data, $\mathcal{G}$ contains subsistence bundles of significant size. It is found that some of the rankings implied by the Ideal Afriat Index are contradicted by the rankings from the Ideal Afriat Marginal Index constructed using another minimum subsistence bundle. Consequently, several countries should perhaps be ranked as equivalent. It is also found that of the 24 countries for which we can make (approximate) true welfare comparisons, 5 (or 20.8 percent) are poor relative to the other countries.

Chapter 6

The third major extension to the RP approach to cross-country welfare comparisons is presented in Chapter 6. The majority of research on cross-country comparisons of welfare are based on per capita expenditure on goods and services; such measures of economic welfare ignore an important commodity, leisure, which a consumer purchases implicitly by not working. In Chapter 6, the joint commodity demand-labour supply framework is used to extend the RP approach to welfare measurement to include the value of leisure time. A multilateral version of the Allen real full income index is shown to be the Leisure-Inclusive Ideal Afriat Marginal Index and this index is proposed for use in leisure-inclusive cross-country comparisons of welfare.

The construction of the Leisure-Inclusive Ideal Afriat Marginal Index requires an estimate of minimum subsistence leisure (that is, time spent eating, sleeping and performing other necessary biological functions). While other authors working on leisure-inclusive welfare measures have needed to make an essentially arbitrary assumption about subsistence leisure, the results from Chapter 5 can be used to nonparametrically estimate this parameter.

For the 1993 ICP data (and using leisure data which has been adjusted for cross-country differences in unemployment rates), an upper-bound estimate of subsistence leisure is 7.7 hours/day. The Leisure-Inclusive Ideal Afriat Marginal Index constructed for this index provides welfare rankings that are very different to those found in the standard leisure-exclusive analysis. In particular, while Luxembourg is ranked first using the Ideal Afriat Index, the fact that the average resident of the U.S. consumes more leisure results in the U.S. heading the leisure-inclusive rankings. Countries with high per capita hours of work slip dramatically in the welfare rankings when leisure is included in the metric: Japan falls from 9th to 19th, and Iceland from 7th to 16th.

Penn 1985 ICP data, β contains subsistence bundles of significant size. It is also found that some of the rankings implied by the Ideal Afriat Index are contradicted by the rankings from the Ideal Afriat Marginal Index constructed using another minimum subsistence bundle. Consequently, several countries should perhaps be ranked as equivalent. It is also found that of the 26 countries for which we can make (approximate) true welfare comparisons, 6 (or 23.1 percent) are poor relative to the other countries.

Chapter 6

The third major extension to the RP approach to cross-country welfare comparisons is presented in Chapter 6. The majority of research on cross-country comparisons of welfare are based on per capita expenditure on goods and services; such measures of economic welfare ignore an important commodity, leisure, which a consumer purchases implicitly by not working. In Chapter 6, the joint commodity demand-labour supply framework is used to extend the RP approach to welfare measurement to include the value of leisure time. A multilateral version of the Allen real full income index is shown to be the Leisure-Inclusive Ideal Afriat Marginal Index and this index is proposed for use in leisure-inclusive cross-country comparisons of welfare.

The construction of the Leisure-Inclusive Ideal Afriat Marginal Index requires an estimate of minimum subsistence leisure (that is, time spent eating, sleeping and performing other necessary biological functions). While other authors working on leisure-inclusive welfare measures have tended to make an essentially arbitrary assumption about subsistence leisure, the results from Chapter 3 can be used to nonparametrically estimate this parameter.

For the 1985 ICP data (and using income data which has been adjusted for cross-country differences in unemployment rates), an (provisional) estimate of subsistence leisure is [illegible]. The Leisure-Inclusive Ideal Afriat Marginal Index constructed for this

Chapter 1

Cross-Country Welfare Comparisons: Theory and Estimation

1.1 Introduction

There are two principal aims of this chapter. First, a selective review of the two main approaches to conducting welfare comparisons is presented.[1] Under the axiomatic approach to index number construction, index numbers are chosen according to their ability to satisfy certain desirable properties or axioms. However, in the construction of these indexes there is no attempt to model the interaction between prices and quantities, and for this reason, indexes constructed under this approach suffer from substitution bias. The economic approach to index number construction, in contrast, is explicitly based on the utility maximising framework and hence interaction between prices and quantities is allowed for.

The second goal of the chapter is to introduce the Revealed Preference (RP) approach to welfare measurement, which is an example of the economic approach to index number construction. The RP approach, and in particular, the multilateral true index which is constructed under this approach, is the focus of the remainder of the thesis. The RP approach can be used to construct both bilateral and multilateral indexes. A multilateral index is defined here as an index which satisfies Fisher's property of circularity - the real income of country i relative to country j should be the same whether the two are

[1]Note that much of the material reviewed in this chapter was originally derived in contexts other than international comparisons (for example, within-country cost-of-living measurement).

compared directly or via an arbitrary intermediate third country. A bilateral index is defined as any index which is not a multilateral index.

The RP approach to constructing bilateral indexes focuses on the estimation of bounds to the Allen welfare index. The Allen index is calculated as a ratio of expenditure functions and hence it is a money metric index of welfare. When preferences across observational units are common, the Allen index exists, however there is an issue of *non-uniqueness* - the welfare comparison between two countries can validly be made using either country's price vector in the calculation. Thus, the base-weighted and current-weighted Allen welfare indexes do not coincide, and the Allen index therefore does not exhibit base-country invariance.

A further problem is that of *indeterminacy* - without complete information on consumer preferences the Allen welfare indexes are unobservable. However, one of the original contributions of this chapter is the development of a framework for constructing (via RP methods) a complete set of bounds to the Allen welfare indexes. First, it is shown that it is possible to construct classical and fixed-weight bounds to the Allen indexes. The indeterminacy of the Allen welfare indexes implies that, conditional on preferences across countries being common, any set of numbers which are contained within the classical and fixed-weight bounds can itself be called a welfare index. This is precisely the definition of a *true welfare index* used in this thesis, namely, it is a set of numbers that satisfies the classical and fixed-weight bounds when preferences are common. In Chapter 2, the RP approach to constructing bilateral true welfare indexes is reviewed in detail. In particular, RP tests of common preferences are outlined and three approaches for constructing tight bounds to the Allen welfare indexes are reviewed.

When preferences are common and homothetic the current-weighted and base-weighted Allen indexes coincide and there exists a unique welfare index. Thus, there is not a problem of non-uniqueness when preferences are homothetic, but the issue of indeterminacy remains. However, it can be shown that with homotheticity, the unique welfare index is bounded by the Paasche and Laspeyres quantity indexes. Further, any set of numbers, the elements of which are contained within the Paasche-Laspeyres (P-L) bounds qualifies as a unique welfare index. Hence, the *unique true welfare index* is defined as a set of numbers that satisfies the P-L bounds.

Finally, it is shown that a unique true welfare index will by construction satisfy circularity, and hence can be called a multilateral true index. The conditions for the existence of a multilateral true index, and how it may be constructed, are further reviewed in Chapter 3, and a particular multilateral true index, the Ideal Afriat Index of Dowrick and Quiggin (1997) is constructed using ICP data.

The structure of the chapter is as follows. In Section 1.2, the basic framework for conducting cross-country comparisons of welfare is discussed. In Section 1.3, the axiomatic approach to index number construction is reviewed, while the economic approach is reviewed in Section 1.4. Three approaches to constructing multilateral indexes of welfare are outlined in Section 1.5. Conclusions are presented in Section 1.6.

1.2 Cross-Country Welfare Comparisons: Preliminaries

In this section, the basic framework for conducting cross-country welfare comparisons is established, and a brief review of the approaches to index number construction used in international comparisons is presented.

1.2.1 Framework for cross-country welfare comparisons

The basic framework for making cross-country comparisons of welfare is as follows. For each of N countries, labeled $i = 1, \dots, N$, it is assumed that there exist observations on the prices (expressed in national currencies) and the quantities (expressed in common units) of K commodities, labeled $l = 1, \dots, K$. Price and quantity vectors in country i are denoted $\mathbf{p}^i$ and $\mathbf{q}^i$, with typical elements p_l^i and q_l^i, respectively. While each good is assumed to be identical in quality worldwide, there are barriers to arbitrage (for example, transport costs or imperfect competition) which prevent prices being equalised internationally.

Relative prices therefore vary across countries, and for this reason, welfare indexes which are constructed using a particular reference price vector (where the typical candidates for this vector have been U.S. prices and weighted averages of prices) are not appropriate for cross-country welfare comparisons. In particular, such *fixed-weight* welfare indexes tend to exhibit the so-called *Gerschenkron effect* (discussed further in Section 1.3.2), namely that a country's relative real income tends to be higher the more the reference prices differ from its own prices. An alternative approach has been to convert national GDP data into a common currency using exchange rates. However, since nontraded goods are generally more labour intensive (and thus relatively cheaper in poorer, labour-abundant countries), exchange rate comparisons tend to overestimate cross-country differences in welfare.

Given the inherent problems in using fixed-weight welfare indexes and exchange-rate adjusted GDP measures, the objective of cross-country welfare comparisons is to find an appropriate set of multilateral index numbers (to be defined below): $\{Q_i,\ i = 1, \dots, N\}$ which imply the bilateral comparisons: $\{Q_{ij} = Q_i/Q_j, \forall i, j = 1, \dots, N\}$.

1.2.2 International Comparisons Programme

The main data source used in this thesis is the International Comparisons Programme (ICP). The ICP provides data on prices and per capita quantities for goods and services that are comparable across countries. The ICP data are highly disaggregated, however more aggregated price and quantity data were used for this thesis. In particular, the data for 1980 cover 60 countries and 18 goods and services, while the data for 1993 cover 23 OECD countries and 19 goods and services (see Appendix A for more detail).

1.2.3 Caveats

There are several caveats that need to be made about the approach to international comparisons used in this thesis. While it is important to qualify the results obtained within this framework in the light of these potential difficulties, one may also note that many of these problems apply to other expenditure-based approaches to international comparisons.

Welfare is based on household expenditure

The framework for international comparisons used here is based on consumer theory, and hence only household consumption data from the ICP are used in the welfare comparisons.[2] Note that household consumption of publicly provided health and education services is included in the definition of household expenditure.

The aim of the thesis is to investigate methods for ranking countries on the basis of living standards, and while investment and government consumption expenditure obviously influence living standards, a thorough investigation of the appropriate theoretical foundations for making international comparisons of investment and government spending was considered outside the scope of the present research.[3] Further, the main aim of the thesis is to extend the RP approach to making cross-country welfare comparisons, and these extensions are valid regardless of exactly what goods are contained in the consumption bundle.

[2]The title of the thesis is therefore somewhat a misnomer, however it is common in the international comparisons literature to equate household expenditure with income.

[3]Dowrick and Quiggin (1994) argue that investment is a claim on future rather than present consumption, and thus include investment in their cross-country welfare comparisons.

Even when the analysis is restricted to household expenditures, there are a number of issues which need to be mentioned. First, it is difficult to accurately measure the quantities and qualities of some services such as education, however the ICP project has made much progress in minimising such difficulties. A further problem is that differences in environmental conditions, for example, are not allowed for in the analysis; an implication of this is that expenditures on heating in cold countries will be measured as an indicator of the well being of the residents. Similar arguments can be made with respect to private expenditures on home security and expenditures addressing health problems related to pollution, for example.[4] A further difficulty with the approach is that the ICP data do not include information on household production nor externalities relating to production and consumption.

The existence of a representative consumer

The framework assumes the existence of a representative or "average" consumer at a national level. This assumption is at odds with intra-country microeconometric evidence that individual preferences do not satisfy the appropriate aggregation conditions (see, for example, Deaton and Muellbauer (1980)) necessary for a representative consumer to capture exactly the aggregate behaviour of heterogeneous households. However, even if individual preferences do not satisfy the appropriate aggregation conditions, aggregate consumption data may still look as if it were generated by a representative consumer if governments redistribute income to maximise a quasi-concave social welfare function (Samuelson (1956, 1964), Varian (1984)).

While such arguments may not satisfy those who work on within-country welfare analysis, the fact is that in the absence of a representative consumer construct, no normative significance can be attached to cross-country comparisons based on per capita expenditures. So the problem of the representative consumer at the national level must be assumed away for the analysis to proceed.

[4]See Eisner (1988) for a discussion of similar problems in the treatment of government expenditures on policing and prisons.

1.2.4 Outline of major approaches to the study of index numbers

An outline of the main approaches to index number construction and welfare comparisons is presented in Figure 1.1.[5] Frisch (1936) distinguishes between two main approaches to index number theory. The axiomatic or test approach, following Fisher (1922), treats observed prices and quantities as independent entities; no attempt is made to model their interaction, although the indexes suggested under this approach are tested for compliance with certain desirable properties or axioms. While the axiomatic approach is useful for identifying properties that an index should possess, it has been criticised for the fact that inter-commodity substitution is not allowed for. The economic approach to index number construction, in contrast to the axiomatic approach, is directly based on utility maximising behaviour and thus attempts to model the interdependencies between prices and quantities which arise from consumer optimising behaviour. The following two sections review these approaches in more detail.

[5] Overviews of the vast literature on index number construction may be found in Diewert (1981, 1987), for example. There are many more methods for multilateral comparisons than those reviewed in this chapter. See, for example, Balk (1996), Diewert (1996) and Hill (1997).

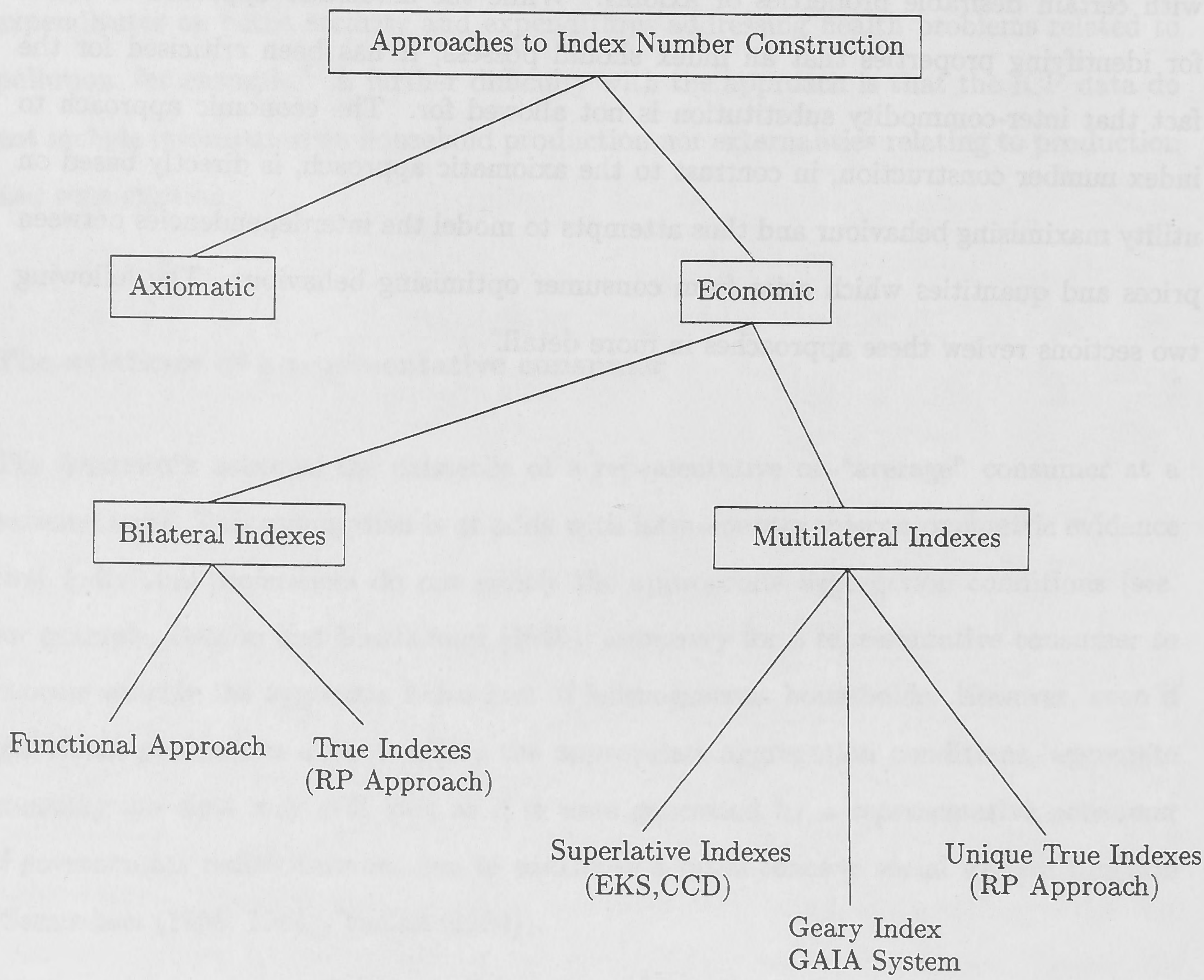

Figure 1.1: Approaches to index number construction

1.3 Axiomatic Approach to Welfare Measurement

In this section, a brief review of the axiomatic approach to welfare measurement is provided. The distinguishing feature of the axiomatic approach is that no attempt is made to model the interaction between prices and quantities.

1.3.1 Some index numbers constructed under axiomatic approach

The simplest measure comparing the welfare of country i with that of country j is an unweighted arithmetic mean of the quantity relatives:[6]

$$Q_{ij} = \frac{1}{K}\sum_{l=1}^{K}\frac{q_l^i}{q_l^j}.$$

Individual quantity observations can have a large impact on the arithmetic mean, and thus it may be more appropriate to use either the geometric or harmonic mean of the quantity relatives, defined respectively:

$$\begin{aligned} Q_{ij} &= \prod_{l=1}^{K}\left(\frac{q_l^i}{q_l^j}\right)^{\frac{1}{K}} \\ Q_{ij} &= \frac{K}{\sum_{l=1}^{K}\frac{q_l^i}{q_l^j}}. \end{aligned}$$

The obvious problem with using these simple index numbers is that all commodities are treated equally; it would be preferable to use weighted index numbers, with the weights based on the budget shares of each commodity.

Weighted arithmetic index numbers

The weighted arithmetic mean of the quantity relatives is:

$$Q_{ij} = \sum_{l=1}^{K}\omega_l^r\frac{q_l^i}{q_l^j},$$

where $\omega_l^r = p_l^r q_l^r / x^r$ is the "reference" budget share for commodity l (and $x^r = \mathbf{p}^r.\mathbf{q}^r$). With base country (country j) budget shares as weights, the weighted arithmetic mean

[6] The following summary draws from Selvanathan and Rao (1994).

of the quantity relatives is the Laspeyres quantity index:

$$Q_{ij}^{L} = \sum_{l=1}^{K} \omega_l^j \frac{q_l^i}{q_l^j} = \frac{\mathbf{p}^j.\mathbf{q}^i}{\mathbf{p}^j.\mathbf{q}^j}. \tag{1.1}$$

Using current country budget shares as weights does not result in any well known index formula, however using the hypothetical budget share resulting from the purchase of current country quantities at base country prices ($\frac{p_l^j q_l^i}{\sum_{l=1}^K p_l^j q_l^i}$) as weights gives the Paasche quantity index:

$$Q_{ij}^{P} = \sum_{l=1}^{K} \frac{q_l^i}{q_l^j} \frac{p_l^j q_l^i}{\sum_{l=1}^K p_l^j q_l^i} = \frac{\mathbf{p}^i.\mathbf{q}^i}{\mathbf{p}^i.\mathbf{q}^j}. \tag{1.2}$$

Other fixed-weight indexes such as the Edgeworth-Marshal and Dobrisch index numbers can be similarly constructed by using appropriated weights in the formula for the weighted arithmetic mean of quantity relatives.

Weighted geometric index numbers

The weighted geometric mean of the quantity relatives (also known as the Cobb-Douglas index) is:

$$Q_{ij} = \prod_{l=1}^{K} \left(\frac{q_l^i}{q_l^j} \right)^{\omega_l^r}.$$

Using ω_l^i or ω_l^j in the above formula does not lead to any recognised indexes, however several index numbers derived according to different specifications of ω_l^r have been proposed:

$$\omega_l^r = \frac{\omega_l^i + \omega_l^j}{2} \qquad \text{(Törnqvist index)}$$

$$\omega_l^r = \frac{\left[\frac{\omega_l^i + \omega_l^j}{2} \omega_l^i \omega_l^j \right]^{1/3}}{\sum_{l=1}^{K} \left[\frac{\omega_l^i + \omega_l^j}{2} \omega_l^i \omega_l^j \right]^{1/3}} \qquad \text{(Theil index)}$$

$$\omega_l^r = \frac{\frac{\omega_l^i \omega_l^j}{\omega_l^i + \omega_l^j}}{\sum_{l=1}^{K} \frac{\omega_l^i \omega_l^j}{\omega_l^i + \omega_l^j}} \qquad \text{(Rao index).}$$

1.3.2 The substitution effect and the Gerschenkron effect

The *consumer substitution effect* is the tendency for a utility maximising consumer to shift spending away from relatively expensive goods towards relatively cheaper goods. In an international comparisons context, the Gerschenkron effect is the finding that fixed-weight indexes tend to overstate the relative real income of a country the more the reference prices differ from its own prices. For example, the ratio of India's real income to that of the U.S. will be higher when U.S. prices are used in the calculation (compared to when India's prices are used). The reason for this is that if the average Indian were actually facing U.S. prices, then he or she would substitute consumption to goods which are relatively cheaper, thus reducing measured expenditure (and hence real income, as defined here). The Gerschenkron effect thus arises because of the substitution effect (and it is also known as *substitution bias*).[7]

As Neary and Gleeson (1997) have argued, substitution bias is not inevitable on theoretical grounds. In fact, they show that the Gerschenkron effect is only a necessary consequence of the representative consumer maximising a homothetic utility function. Further, Hill (2000) notes that the *producer substitution effect*, which is the tendency for profit maximising producers to substitute production away from relatively cheaper goods towards relatively more expensive goods, acts in the opposite direction to the consumer substitution effect. However, Hill (2000) argues that in practice, at least at the level of GDP, the consumer substitution effect always dominates the producer substitution effect.

The conclusion is that for a given data set, the empirical relevance of the Gerschenkron effect should be investigated, rather than assumed. Neary and Gleeson (1997) devised a method for evaluating the magnitude of the Gerschenkron effect which involves regressing the Laspeyres/Paasche ratio on a measure of price dispersion between the two countries (the sum of squared deviations between the prices of individual goods).[8] The authors find a positive significant coefficient in the regression for 1970 ICP data on 16 countries and 11 categories of consumption, and thus argue that the Gerschenkron effect is present in these data. As discussed further below, Dowrick and Quiggin (1997) and Hill (2000) have devised other methods for detecting substitution bias in international comparisons and provide evidence of this bias in more recent ICP data sets.

[7]Gerschenkron (1951) observed the dependence of the growth rate of Soviet machinery output between 1927-28 and 1938, and of U.S. machinery output between 1899 and 1939, on the choice of reference price vector.

[8]A consequence of the Gerschenkron effect is that the Laspeyres index will exceed the Paasche index.

1.3.3 Substitution bias, income dispersion and convergence

The presence of substitution bias leads to two important implications for international comparisons. First, fixed-weight real per capita income measures tend to underestimate (overestimate) the dispersion of income levels when a relatively rich (poor) country's prices are used to value income. Dowrick and Quiggin (1997) found for 1980 and 1990 ICP data that the variance of log per capita GDP across OECD countries tended to be higher when real GDP was constructed using the prices of poorer countries such as Greece and Portugal, and lower when measured using Canadian and U.S. prices.

However, they also found an inconsistency in this pattern, namely that the dispersion of GDP constructed using Spanish prices in 1980 was low, compared to that found using the prices of other poor countries. The authors reasoned that the magnitude of substitution bias depends on both the reference country's relative income level *and* on the dissimilarity of prices. In 1980, Spanish prices were remarkably close to average international prices, and this explains the above finding.

For 1985 ICP data, Hill (2000) found a similar pattern with Gini coefficients, namely that they tended to be lower (higher) when calculated using the prices of rich (poor) countries.[9] However, an inconsistency in this pattern was that the lowest Gini coefficients were calculated using the price vectors of the relatively poor Caribbean countries. As above, a possible explanation for this finding is that the prices of the Caribbean countries were very similar to the international average. If this is the explanation for the finding (rather than a data quality issue, for example), then it suggests that the Gini coefficient might be very sensitive to the similarity of the reference country's prices to the international average, compared with other measures of inequality such as variance of log per capita income.

The second implication of the presence of substitution bias is that the use of fixed-weight indexes will tend to confuse the measurement of the rate of convergence in per capita income over time with convergence in prices (Nuxoll (1994) and Dowrick and Quiggin (1997)). Dowrick and Quiggin (1997) show this by calculating σ convergence - the measured change in the variance (σ^2) of log per capita GDP between 1980 and

[9] The Gini coefficient is a measure of the equality of the income distribution; a Gini of zero indicates absolute equality, while a Gini of one indicates absolute inequality.

1990 - using the price vectors for each of the countries in the data sets. They found evidence of σ convergence (that is, the variance of log GDP falls between 1980 and 1990) for all price vectors except that of Luxembourg (for which there was very weak σ divergence). In addition, Dowrick and Quiggin (1997) found that rich country prices tend to give estimates of weak σ convergence, while poor country prices imply that the rate of convergence is much higher.[10]

The reason for this finding is related to the fact that prices tended to converge over the 1980s. This was due to both increases in trade and also diminishing variation in the relative price of nontradeables (which are primarily comprised of services) as a result of real wage convergence. The convergence of prices means that the level of substitution bias was lower in 1990, compared with 1980. Hence, it is apparent that if we use a rich country's prices, the variance of log GDP in both 1980 and 1990 will be biased downwards, but the degree of downward bias will be less in 1990 (because substitution bias is less). Thus, if there has been σ convergence (that is, the variance of log GDP is higher in 1980 compared with 1990), we will underestimate this convergence when rich country prices are used in the calculation. The converse is true when poor country prices are used.

1.3.4 Index numbers constructed from fixed-weight indexes

It has been argued that index numbers constructed as functions of fixed-weight indexes such as the Laspeyres and Paasche indexes will be subject to less substitution bias, and hence can be considered an improvement. Such index numbers are termed *empirical* index numbers, and the best-known bilateral empirical index number is that suggested by Fisher:

$$Q_{ij}^{F} = \sqrt{Q_{ij}^{L} Q_{ij}^{P}}.$$

This index is known as the Ideal Fisher Index, in acknowledgment of its ability to satisfy a number of useful axioms, as discussed below. While an empirical index number such as the Ideal Fisher Index is not subject to substitution bias, the correction for the substitution effect is of a *statistical* nature and does not involve economic theory. This contrasts with the economic approach discussed below, where substitution bias is corrected for

[10] Once again, with the exception of Spain.

with explicit reference to utility maximising behaviour.

1.3.5 Desirable properties of an index number

Index numbers constructed under the axiomatic approach are constructed without specific reference to utility maximisation. The question then is how may one chose between the different quantity indexes constructed under this approach? Fisher (1922) suggested a number of tests that may be used to select an appropriate index number formula (hence the axiomatic approach to index number construction is also known as the test approach).[11]

Four axioms are of particular relevance to multilateral comparisons:

1. *Base-country invariance:*[12] It is desirable that in comparing the welfare of two countries, indexes of real income should not be sensitive to the choice of base-country: $Q_{ij} = 1/Q_{ji}$.
2. *Circularity:* The real income of country i relative to country j should be the same whether the two are compared directly or via an arbitrary intermediate third country k: $Q_{ij} = Q_{ik}Q_{kj}$.[13]
3. *Characteristicity* or *Independence of Irrelevant Countries:* The comparison between two countries should only depend on variables which characterise them and not on variables characteristic of other countries. Thus, country i's real income relative to that of country j should be unaffected by changes in the real income of country k.
4. *Matrix consistency:* The usefulness of a set of real income indexes is enhanced if they can be consistently disaggregated by commodity as well as by country.

Having established the criteria deemed desirable for cross-country welfare analysis, the question is: how well do the index numbers constructed above satisfy these criteria?

[11] Diewert (1992) provides a comprehensive summary of these axioms.

[12] This is also known as the point reversal test.

[13] Note that the Circularity test originally stated by Fisher was: $Q_{ik}Q_{kj}Q_{ji} = 1$. Circularity implies: $Q_{ii} = 1$ (this is known as the Identity test) and $Q_{ij}Q_{ji} = 1$ (base-country invariance). Further, circularity can be seen to be equivalent to $Q_{ij} = Q_{ik}Q_{kj}$, which Fisher originally referred to as the Chain test. Thus the Chain test is equivalent to circularity, and these tests are used interchangeably in the literature. It is apparent that circularity is more stringent than base-country invariance and in fact implies it (replace j with i in the above formula). The property of circularity is sometimes referred to as *transitivity* (see, for example, Drechsler (1973)), however since this may lead to confusion with transitivity of the underlying preferences (which is an ordinal concept), the term is not used here.

The fixed-weight Paasche and Laspeyres indexes do not satisfy base-country invariance and while the Ideal Fisher Index does satisfy this property, it fails the stricter test of circularity. The inability of the Fisher Index to satisfy circularity makes it of limited use in multilateral comparisons of welfare, and has led to multilateralisations of this index which do satisfy circularity (see Section 1.5.2).

1.4 Economic approach to Welfare Measurement

While the test approach provides a framework for choosing between different empirical index numbers, ever since Frisch (1936) this approach has been subject to criticism on a number of grounds.[14] At a practical level, there is a problem that the tests often turn out to be mutually inconsistent (see Eichhorn and Voeller (1976) for details). At a theoretical level, the test approach does not require indexes to have a basis in economic theory and hence fixed-weight indexes constructed under this approach are subject to substitution bias. Finally, at a conceptual level, empirical index numbers are subject to the criticism of Afriat (1977) that they provide "answers without questions": since there is no reference to economic theory, it is not clear what a real income index constructed under the axiomatic approach is in fact measuring.

The economic approach to index numbers avoids these difficulties by explicitly starting from utility maximising behaviour. Thus, in the context of international comparisons, the data are assumed to be generated by the utility maximising behaviour of representative consumers in each country, each sharing identical tastes. The difference between the RP approach to welfare measurement and the other economic approaches to index number construction is that in the RP approach, this assumption of common preferences across countries is tested for using revealed preference relations.

Neary (2000) has argued that there are two distinct questions which must be answered in using the economic approach to choosing an appropriate index for multilateral comparisons. First, exactly what index is to be measured? Second, given a particular choice of welfare index, which by nature is unobservable, how can it be best approximated? Since the economic approach postulates the existence of indexes of welfare, the various indexes constructed under this approach may be judged by how closely they approximate these indexes.

1.4.1 A selective review of consumer theory

As the economic approach to index number construction and welfare measurement is explicitly derived from utility maximisation framework, it is useful to first review some

[14]This summary is from Neary (2000).

relevant results from consumer theory.

The consumer's problem of choosing the most preferred consumption bundle given $\mathbf{p}$ and total expenditure x is the following *utility maximisation problem*:

$$\max_{\mathbf{q}\geq 0} u(\mathbf{q}) \quad \text{subject to}$$
$$\mathbf{p}.\mathbf{q} \leq x,$$

where $u(\mathbf{q})$ is the utility function. The utility maximisation problem thus defines the *Marshallian* demand function, $\mathbf{q}(x, \mathbf{p})$; this is the rule which assigns the set of optimal consumption vectors to each expenditure-price situation.

The dual to the utility maximisation problem is the *expenditure minimisation problem*:

$$\min_{\mathbf{q}\geq 0} \mathbf{p}.\mathbf{q} \quad \text{subject to}$$
$$u(\mathbf{q}) \geq U,$$

where U is a particular level of utility. This problem thus defines the expenditure-minimising demand function, denoted $\mathbf{h}(U, \mathbf{p})$, which is also known as the *Hicksian* or *compensated* demand function.

Note that the same U and x are used in both problems: the U in the expenditure minimisation problem is the maximum attainable in the utility maximisation problem and since utility maximisation and expenditure minimisation must imply the same choice, the outlay in the original problem must be the expenditure minimum in the dual problem. Since the solutions to both problems coincide, it must be that $\mathbf{q}(x, \mathbf{p}) = \mathbf{h}(U, \mathbf{p})$.

Each of these solutions can be substituted back into their respective problems to give the maximum attainable utility and minimum attainable expenditure, respectively. The maximum attainable utility given $\mathbf{p}$ and x is given by the *indirect utility function*, which is defined:

$$\psi(x, \mathbf{p}^i) = \max_{\mathbf{q}}[u(\mathbf{q}) : \mathbf{p}^i.\mathbf{q} = x].$$

The minimum expenditure associated with attaining U at prices $\mathbf{p}$ is given by the *expenditure function*, which is defined:

$$e(U, \mathbf{p}^i) = \min_{\mathbf{q}}[\mathbf{p}^i.\mathbf{q} : u(\mathbf{q}) = U].$$

Thus, $e(U, \mathbf{p}^i)$ can be thought of as the minimum cost of reaching the indifference curve labeled U, when prices are $\mathbf{p}^i$.

Money metric utility

Demand theory is completely ordinal in nature - there is no unique cardinal representation of utility. Indifference curves are unaffected by monotonic increasing transformations of $u(\mathbf{q})$, so that by moving to a transformation $f[u(\mathbf{q})]$, we simply re-label the identical indifference curves. For example, if we know preferences are represented by the utility function $u(\mathbf{q})$, then the indifference curve for a particular bundle $\mathbf{q}^i$ will be labeled $U^i = u(\mathbf{q}^i)$. However, preferences may equivalently be represented by $f[u(\mathbf{q})]$, and for this transform, the indifference curve for the bundle $\mathbf{q}^i$ will be labeled $f^{-1}[U^i]$, where $f^{-1}[]$ is the inverse function of $f[]$.

The implication of the ordinality of demand is that the utility function (even if it is known, which is not the case in practice) cannot be directly used for computing ratios of the welfare of different countries. The reason for this is that the transformations of the utility functions will result in different welfare ratios, and since all transforms of the utility function are equally valid representations of preferences, there is no reason to choose one over any other.

For example, suppose preferences are represented by the following Stone-Geary utility function, $u(\mathbf{q}) = (q_1 - 0.5)^{1/2}(q_2 - 0.75)^{1/2}$ and we are wanting to compare the welfare of two countries with the following prices and demand data (note that the demand data have been generated from the utility function): $\mathbf{p}^1 = \{1, 1\}$, $\mathbf{q}^1 = \{1.875, 2.125\}$ and $\mathbf{p}^2 = \{2, 1\}$, $\mathbf{q}^2 = \{2.563, 4.875\}$. The welfare of country 1 compared with 2 can be calculated as: $U^1/U^2 = \frac{(q_1^1-0.5)^{1/2}(q_2^1-0.75)^{1/2}}{(q_1^2-0.5)^{1/2}(q_2^2-0.75)^{1/2}} = 1.375/2.917 = 0.471$. However, it can also validly be calculated using the logarithm of the utility function $U^* = \log[u(\mathbf{q})]$ as: $U^{*1}/U^{*2} = \log 1.375/\log 2.917 = 0.297$.

The above example illustrates how the welfare comparison between two countries is dependent on which monotonic transformation of the utility function is used. However, it is common practice to remove this source of non-uniqueness in welfare comparisons by using certain cardinalisations of utility. One such cardinalisation is Samuelson's (1974) *money metric utility function*; the cost of attaining utility level U at reference prices $\mathbf{p}^r$

is $e(U, \mathbf{p}^r)$; with prices fixed, the expenditure function is a monotonic increasing function of U and hence can be considered a measure of utility in its own right.[15] The use of the money metric utility function is equivalent to "assuming away" all the possible monotonic transformations of the utility function. However, as shown in the next sub section, money metric utility introduces another source of non-uniqueness which is related to the choice of the reference price vector in the welfare comparison.

Homothetic preferences

The final part of this review of consumer theory relates to a particular form of preferences which is used throughout this thesis. Formally, preferences are homothetic if and only if utility is a monotone increasing function of a function that is homogeneous of degree 1:[16]

$$U = f[u(\mathbf{q})]; \qquad u(\theta\mathbf{q}) = \theta u(\mathbf{q}),$$

where θ is an arbitrary positive scalar. Since consumer behaviour is invariant with respect to monotonic transformations of the utility function, we can choose $u(\mathbf{q})$ itself as the representation of preferences (i.e. we can ignore $f[]$).

The property of homotheticity implies that the slope of the indifference curves are constant along a ray from the origin. Thus, the income expansion paths are rays from the origin and the budget shares are invariant to changes in income (the income elasticities are constrained to one). However, it should be noted that these properties hold when utility is an increasing monotonic transformation of a utility function which is homogeneous *of any degree*, and are not restricted to homotheticity.

When $u(\mathbf{q})$ is linearly homogeneous, a doubling of U will be accomplished by doubling $\mathbf{q}$ and, hence, by doubling costs. Therefore, the expenditure function will be linearly homogeneous in U:

$$e(U, \mathbf{p}^i) = Ue(1, \mathbf{p}^i) = Ub(\mathbf{p}^i),$$

where $b(\mathbf{p}^i) = e(1, \mathbf{p}^i)$ is linearly homogeneous and concave in $\mathbf{p}^i$. Under homotheticity,

[15] The money metric utility function is also known as the "minimum income function", and the "direct compensation function". An alternative definition used by Varian (1982) is: $m(\mathbf{q}, \mathbf{p}^r) \equiv e(u(\mathbf{q}), \mathbf{p}^r)$. This definition of money metric utility is used further in Chapter 2.

[16] See Appendix B for a further discussion of the property of homogeneity.

the indirect utility function can therefore be calculated as:

$$\psi(x, \mathbf{p}^i) = e(U, \mathbf{p}^i)/b(\mathbf{p}^i) = x/b(\mathbf{p}^i) = \psi(\mathbf{p}^i)x,$$

where $\psi(\mathbf{p}^i) = b(\mathbf{p}^i)^{-1}$.

1.4.2 Three welfare indexes

There are three indexes of welfare which have been proposed. Following Deaton and Muellbauer (1980), *welfare* indexes are distinguished here from *quantity* indexes by the fact that in order to calculate the former, complete knowledge of the form of preferences is required, while quantity indexes can be calculated with knowledge of the actual demands (and the prices).[17] It is important to note that while some authors refer to the following three welfare indexes as *true* welfare indexes, the adjective "true" is not used here in this context since it has a precise definition which is introduced later in this chapter.

The Allen welfare index

The Allen index comparing utility levels U^i and U^j at reference prices $\mathbf{p}^r$ is:

$$Q_{ij}^{A,r} = e(U^i, \mathbf{p}^r)/e(U^j, \mathbf{p}^r).$$

The Allen index therefore gives the fraction of the cost of attaining country j's utility level required to attain country i's utility level at the reference prices. If, for example, $Q_{ij}^{A,r} = 2$, then it costs twice as much at prices $\mathbf{p}^r$ to attain i's utility level as it does to attain j's utility level and thus, under the money metric concept of utility, consumer i is said to be twice as well off as consumer j. Since by definition the expenditure function gives the minimum cost of attaining a given utility level when facing given prices, the Allen index allows for inter-commodity substitution.

The natural candidates for $\mathbf{p}^r$ are $\mathbf{p}^j$ and $\mathbf{p}^i$, giving base-weighted and current-weighted

[17] Similarly, the calculation of *cost-of-living* indexes require information on preference, while *price* indexes again can be calculated with information only on actual quantities demanded and the prices. For reasons of brevity, welfare indexes defined using the indirect utility function (see, for example, Lloyd (1979)) are not reviewed here.

welfare indexes, which are known as the Laspeyres-Allen and Paasche-Allen welfare indexes, respectively:

$$Q_{ij}^{LA} = e(U^i, \mathbf{p}^j)/e(U^j, \mathbf{p}^j)$$
$$Q_{ij}^{PA} = e(U^i, \mathbf{p}^i)/e(U^j, \mathbf{p}^i).$$

The Konüs welfare index

The cost-of-living index, which is due to Konüs, compares the cost of achieving a particular level of utility, U^r, given different prices:

$$P_{ij}^{K,r} = e(U^r, \mathbf{p}^i)/e(U^r, \mathbf{p}^j).$$

The Laspeyres-Konüs and Paasche-Konüs cost-of-living indexes are defined, respectively, as:

$$P_{ij}^{LK} = e(U^j, \mathbf{p}^i)/x^j$$
$$P_{ij}^{PK} = x^i/e(U^i, \mathbf{p}^j),$$

where $x^i = e(U^i, \mathbf{p}^i)$.

It is desirable that the product of a price and welfare index equal the ratio of actual expenditures between the two countries (this property is known as the *weak factor reversal test*). For the Allen welfare index, in the absence of homothetic preference, this property will only hold when the quantity index is base-weighted and the price index is current-weighted (and vice-versa) i.e., $Q_{ij}^{LA} P_{ij}^{PK} = x^i/x^j$ and $Q_{ij}^{PA} P_{ij}^{LK} = x^i/x^j$.

An alternative to the Allen welfare index is the Konüs welfare index, which satisfies weak factor reversal by construction:

$$Q_{ij}^{K,r} = \frac{x^i}{x^j P_{ij}^{K,r}}.$$

The Laspeyres-Konüs and Paasche-Konüs welfare indexes are defined, respectively:

$$Q_{ij}^{LK} = \frac{x^i}{x^j P_{ij}^{LK}} = \frac{x^i}{e(U^j, \mathbf{p}^i)}$$
$$Q_{ij}^{PK} = \frac{x^i}{x^j P_{ij}^{PK}} = \frac{e(U^i, \mathbf{p}^j)}{x^j},$$

and it is immediately apparent that $Q_{ij}^{LK} = Q_{ij}^{PA}$ and $Q_{ij}^{PK} = Q_{ij}^{LA}$.

The Malmquist welfare index

The Allen and Konüs indexes are both constructed using an approach where each indifference curve is in effect labeled by the expenditure required to reach it. An alternative approach is label indifference curves by their relative distance from the origin along some reference quantity bundle. The Malmquist *quantity* index is defined:[18]

$$Q_{ij}^{MQ,r} = \frac{d(U^r, \mathbf{q}^i)}{d(U^r, \mathbf{q}^j)},$$

where the *distance function*, $d(U^r, \mathbf{q}^i) = \max_\delta\{\delta : u(\mathbf{q}^i/\delta) \geq U^r\}$. The distance function therefore measures the scalar by which $\mathbf{q}^i$ must be divided by so that $u[\mathbf{q}^i/d(U^r, \mathbf{q}^i)] = U^r$. Note that by definition, $d(U^i, \mathbf{q}^i) = 1$, for all i.

The Laspeyres-Malmquist and Paasche-Malmquist quantity indexes are defined, respectively:

$$Q_{ij}^{LMQ} = \frac{d(U^j, \mathbf{q}^i)}{d(U^j, \mathbf{q}^j)} = \frac{d(U^j, \mathbf{q}^i)}{1}$$

$$Q_{ij}^{PMQ} = \frac{d(U^i, \mathbf{q}^i)}{d(U^i, \mathbf{q}^j)} = \frac{1}{d(U^i, \mathbf{q}^j)}.$$

While it is possible to use Q_{ij}^{LMQ} and Q_{ij}^{PMQ} as welfare indexes, Deaton and Muellbauer (1980, p.182) define the Laspeyres-Malmquist and Paasche-Malmquist *welfare* indexes respectively:

$$Q_{ij}^{LM} = \frac{d(U^j, \mathbf{q}^j)}{d(U^i, \mathbf{q}^j)} = \frac{1}{d(U^i, \mathbf{q}^j)}$$

$$Q_{ij}^{PM} = \frac{d(U^j, \mathbf{q}^i)}{d(U^i, \mathbf{q}^i)} = \frac{d(U^j, \mathbf{q}^i)}{1}.$$

It is apparent that:

$$Q_{ij}^{LM} = Q_{ij}^{PMQ}$$

$$Q_{ij}^{PM} = Q_{ij}^{LMQ}.$$

An advantage of the Malmquist welfare indexes is that they are defined independently of the existence of prices - this may be particularly useful in situations of rationing or when the budget constraint is nonlinear. Note that while the Malmquist index is not

[18]The notational convention adopted above for distinguishing between quantity and welfare indexes does not hold with the Malmquist index, since the unobservable utility function appears in both indexes.

constructed using money metric utility, its construction still involves using a particular cardinalisation of utility and therefore all other possible positive monotonic transformations of the utility function are ignored. Finally, it can be noted (see Diewert (1981)) that:

$$Q_{ij}^{LM} \geq Q_{ij}^{LA}$$

$$Q_{ij}^{PM} \leq Q_{ij}^{PA}.$$

Calculating the welfare indexes - Stone-Geary example

The calculation of the welfare indexes is now illustrated using the above two-good two-country example where preferences are Stone-Geary.[19] The general Stone-Geary utility function is $u(\mathbf{q}) = \prod_{l=1}^{K}(q_l - \gamma_l)^{\beta_l}$, $\sum_l \beta_l = 1$, where γ_l is minimum subsistence consumption of good l. The Marshallian demand for good l is $p_l q_l = p_l \gamma_l + \beta_l(x - \sum_l p_l \gamma_l)$. Substituting the Marshallian demand functions into the utility function gives the indirect utility function: $\psi(x, \mathbf{p}) = (x - \sum_l p_l \gamma_l)/\beta_0 \prod_l p_l^{\beta_l}$, where $\beta_0 = 1/\prod_l \beta_l^{\beta_l}$. Inverting this gives the expenditure function: $e(U, \mathbf{p}) = \beta_0 \prod_l p_l^{\beta_l} U + \sum_l p_l \gamma_l$.

For the two-good two-country example, the Allen welfare indexes are:

$$Q_{12}^{LA} = \frac{\beta_0 p_1^{2\beta_1} p_2^{2\beta_2} U^1 + (p_1^2\gamma_1 + p_2^2\gamma_2)}{x^2} = 0.5639$$

$$Q_{12}^{PA} = \frac{x^1}{\beta_0 p_1^{1\beta_1} p_2^{1\beta_2} U^2 + (p_1^1\gamma_1 + p_2^1\gamma_2)} = 0.5646.$$

The Konüs welfare indexes are:

$$Q_{12}^{LK} = \frac{x^1}{\beta_0 p_1^{1\beta_1} p_2^{1\beta_2} U^2 + (p_1^1\gamma_1 + p_2^1\gamma_2)} = 0.5646$$

$$Q_{12}^{PK} = \frac{\beta_0 p_1^{2\beta_1} p_2^{2\beta_2} U^1 + (p_1^2\gamma_1 + p_2^2\gamma_2)}{x^2} = 0.5639.$$

The Malmquist welfare indexes are:

$$Q_{12}^{LM} = 0.5640$$

$$Q_{12}^{PM} = 0.5644.$$

[19] The following exercise is an example of the functional approach to index number construction (which itself comes under the economic approach). The functional approach involves estimating the parameters of utility or expenditure functions (usually via estimation of the associated demand system) and then using those parameters to construct the various welfare indexes.

1.4.3 Classical and fixed-weight bounds to Allen welfare indexes

The majority of empirical work on constructing welfare indexes has focused on the Allen indexes Q_{ij}^{LA} and Q_{ij}^{PA} and therefore these are the focus of the following discussion. In the above example, the Allen indexes were calculated for a particular form of the Stone-Geary utility function; in practice, we do not know the form of the utility function and hence the Allen indexes cannot be directly calculated. However, given that the Allen indexes exist, it is possible to construct *bounds* to these indexes using only the demand data (that is, the price and quantity data). The condition for the Allen indexes to exist is now stated.

Result 1.1. Existence of the Allen Welfare Indexes - Various:[20] *The Allen Welfare Indexes Q_{ij}^{LA} and Q_{ij}^{PA} exist if and only if the demand data can be rationalised by a representative consumer.*

A formal proof of this statement is not given here. It is noted that if the consumers didn't share common preferences (and hence maximise a common utility function), then their expenditure functions would not be of the same form and the Allen indexes are defined for common expenditure functions.

In Chapter 2, the different approaches to testing whether a given data set is consistent with the null hypothesis of common preferences are reviewed. In Proposition 2.1 it is shown that a given set of data can be rationalised by a utility function if the Generalised Axiom of Revealed Preference (GARP) is satisfied. For the case of $N = 2$, the test of common preferences reduces to the following condition (see Chapter 2 for further details):

Proposition 1.1. Revealed Preference Test of Common Preferences with $N = 2$ - Various: *With $N = 2$, we do not reject the hypothesis that preferences across country i and country j are common if and only if the data satisfy the following condition:*[21]

$$\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}^j \ \textit{implies} \ \mathbf{p}^j.\mathbf{q}^j \leq \mathbf{p}^j.\mathbf{q}^i,$$

[20] The convention in this thesis is to attribute all propositions and results to particular authors (or to "various"). If no attribution is given directly or implied by the text then the proposition or result is original.

[21] Note that this condition is *not* the Weak Axiom of Revealed Preference (WARP). As shown in Chapter 2, WARP requires that there be a unique demand bundle at each budget, while the condition shown here allows for multi-valued demand functions.

which is equivalent to:

$$Q^L_{ji} \leq 1 \text{ implies } Q^P_{ji} \leq 1.$$

PROOF: See proof to Proposition 2.1.

Given that the data do not reject the null hypothesis of common preferences then it is possible to construct bounds to the unobservable welfare indexes. In particular, *fixed-weight* and *classical* bounds can be constructed. To this author's knowledge, previous discussion on classical bounds to indexes has only been in the context of cost-of-living indexes (see, for example Varian (1982) and Manser and McDonald (1988)). The following presentation is an adaptation of this previous work to the construction of classical bounds to welfare indexes.[22]

The construction of the classical and fixed-weight bounds to the Allen indexes involves finding bounds to the (unobservable) expenditure function, $e(U^i, \mathbf{p}^j)$. The upper bound to this function is calculated by noting that since consuming the bundle $\mathbf{q}^i$ is one way of achieving U^i, but not necessarily the cheapest when prices are $\mathbf{p}^j$, it must therefore be that $\mathbf{p}^j.\mathbf{q}^i \geq e(U^i, \mathbf{p}^j)$.[23]

The construction of a lower bound to $e(U^i, \mathbf{p}^j)$ involves the concept of the *classical pseudo-expenditure function*, which is defined:

$$e_C(U^i, \mathbf{p}^j) = \min_{\mathbf{q}} \mathbf{p}^j.\mathbf{q} \text{ subject to } \mathbf{p}^i.\mathbf{q} = x^i.$$

The function $e_C(U^i, \mathbf{p}^j)$ is thus calculated as the cost (at prices $\mathbf{p}^j$) of the cheapest bundle that is just affordable at prices $\mathbf{p}^i$ and income x^i. Consumption of this bundle will typically give less utility than U^i, and thus cost less. Hence, by construction, $e_C(U^i, \mathbf{p}^j)$ must be (weakly) less than $e(U^i, \mathbf{p}^j)$. Note that the classical pseudo expenditure function

[22] There have been differing definitions of the classical bounds to the cost-of-living index used in the literature. For example, Varian (1982, p.967), and the NONPAR software which accompanies that article, uses a definition of the classical bound which is the inspiration for the definition employed here. Other authors have made use of the bounds provided by Lerner (1935) namely that the cost-of-living index (being some weighted average of price changes) must lie somewhere between the maximum and minimum of the ratio of the prices:

$$\min_l \{p^i_l/p^j_l : l = 1, \dots k\} \leq P^{K,r}_{ij} \leq \max_l \{p^i_l/p^j_l : l = 1, \dots k\}.$$

Pollak (1971) improved these bounds by noting that the Laspeyres index bounds the base-weighted cost-of-living index from above. Hence authors such as Manser and McDonald (1988) and Blundell, Browning, and Crawford (1998) refer to the classical bounds to the cost-of-living index as the Laspeyres price index from above and the minimum price relative from below.

[23] The reflexivity of the preference relation guarantees this.

$e_C(U^i, \mathbf{p}^j)$ is only introduced here for heuristic purposes (to show the construction of the classical bounds to the bilateral true welfare indexes). It is called a pseudo expenditure function since, unlike the expenditure function, it can be calculated without knowledge of the form of preferences. That is, U^i obviously does not figure in the calculation of $e_C(U^i, \mathbf{p}^j)$, however it can be said that a consumer spending this amount when facing prices $\mathbf{p}^j$ could not achieve a level of utility higher than U^i (and typically will achieve a lower level of utility, unless there is a corner solution).

The fixed-weight and classical approximations to $e(U^i, \mathbf{p}^j)$ are shown in Figure 1.2 (where the diagram has been constructed under the assumption that $p_2 = 1$, and hence expenditure can be measured along the vertical axis).

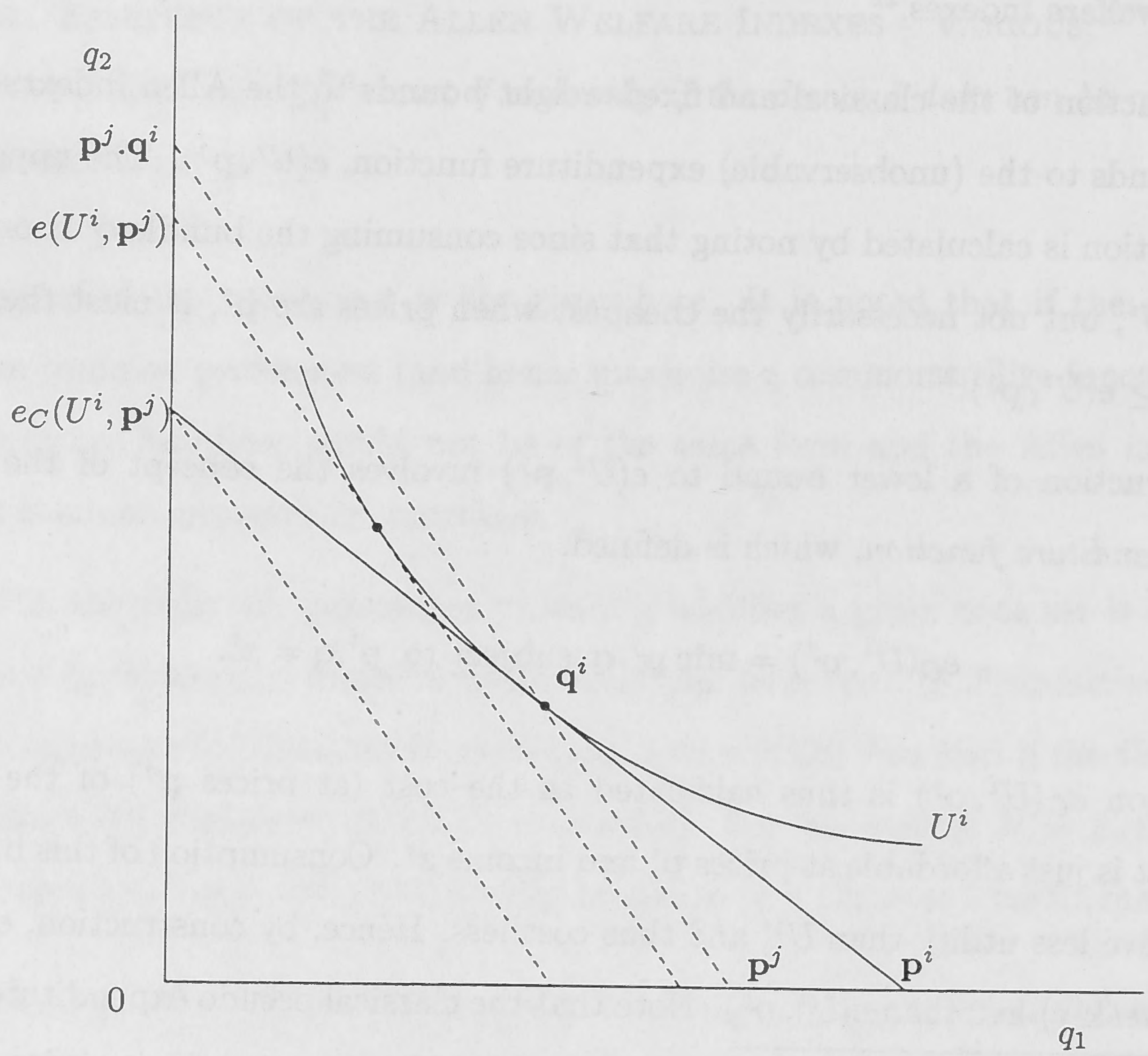

Figure 1.2: The fixed-weight and classical approximations to $e(U^i, \mathbf{p}^j)$

The fixed-weight and classical bounds to the expenditure function can therefore be stated as:

(1.3) $$e_C(U^i, \mathbf{p}^j) \leq e(U^i, \mathbf{p}^j) \leq \mathbf{p}^j.\mathbf{q}^i.$$

To further illustrate the construction of the classical and fixed-weight bounds to the expenditure function, we can consider particular examples of preferences and price vectors. In Figure 1.3, Leontief preferences (where there is no substitution between goods in response to relative price changes) are illustrated and in this case, it is apparent that $e_C(U^i, \mathbf{p}^j) \leq e(U^i, \mathbf{p}^j) = \mathbf{p}^j.\mathbf{q}^i$.

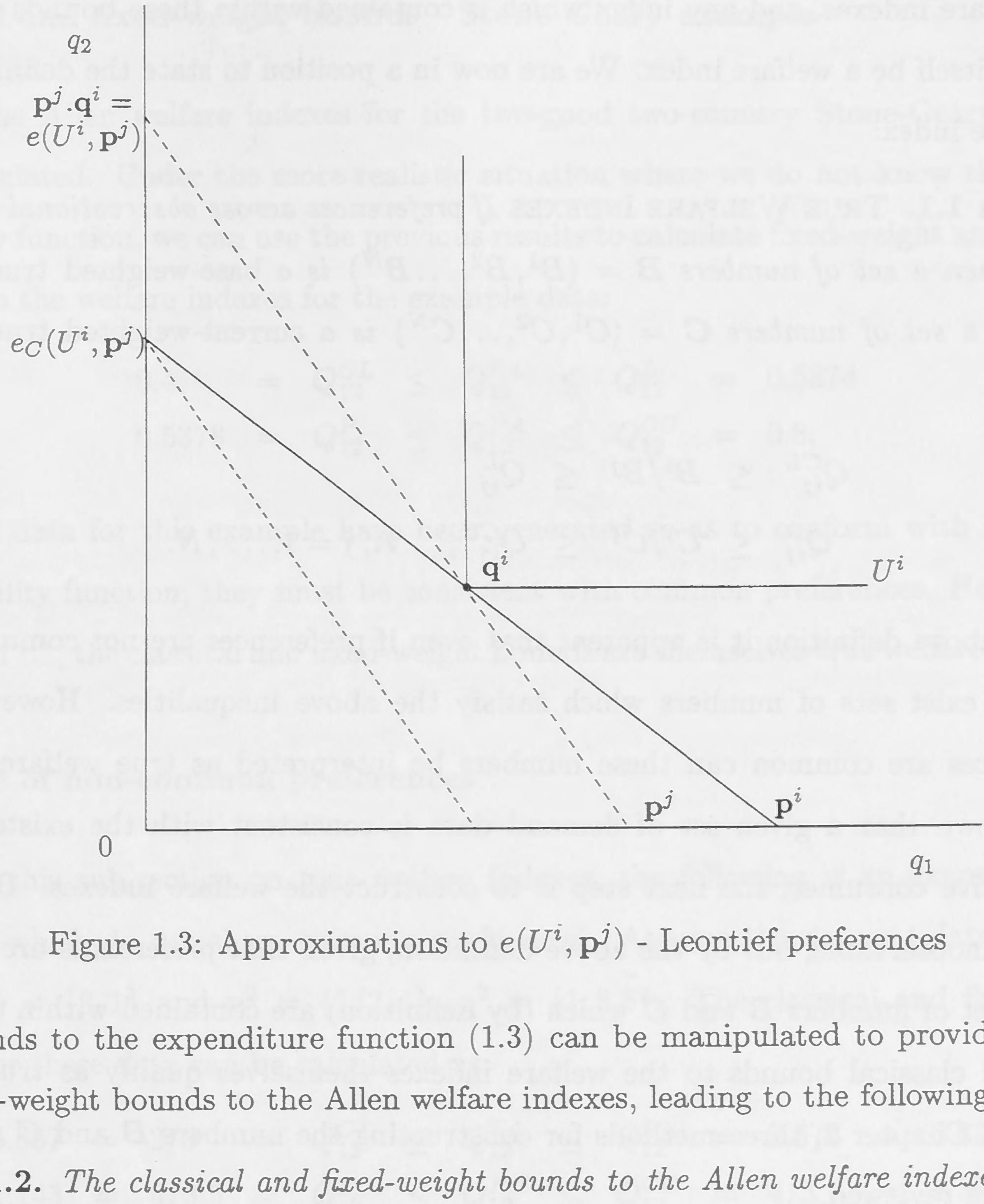

Figure 1.3: Approximations to $e(U^i, \mathbf{p}^j)$ - Leontief preferences

The bounds to the expenditure function (1.3) can be manipulated to provide classical and fixed-weight bounds to the Allen welfare indexes, leading to the following result:

Result 1.2. *The classical and fixed-weight bounds to the Allen welfare indexes are:*

$$Q_{ij}^{CL} \leq Q_{ij}^{LA} \leq Q_{ij}^{L}$$
$$Q_{ij}^{P} \leq Q_{ij}^{PA} \leq Q_{ij}^{CU},$$

where Q_{ij}^{L} and Q_{ij}^{P} are, respectively, the fixed-weight upper bound to Q_{ij}^{LA} and fixed-weight lower bound to Q_{ij}^{PA}, and the classical lower bound to Q_{ij}^{LA} and classical upper bound to Q_{ij}^{PA} are respectively:

$$Q_{ij}^{CL} = e_C(U^i, \mathbf{p}^j)/x^j$$
$$Q_{ij}^{CU} = x^i/e_C(U^j, \mathbf{p}^i).$$

In summary, when preferences across observational units (e.g., countries) are common, then the Allen welfare indexes exist. Since the exact form of the utility function is unknown, we are not able to directly construct these indexes and hence there is a problem of *indeterminacy*. However, we are able to construct the fixed-weight and classical bounds to the welfare indexes, and any index which is contained within these bounds must, by definition, itself be a welfare index. We are now in a position to state the definition of a true welfare index:

Definition 1.1. TRUE WELFARE INDEXES *If preferences across observational units are common then a set of numbers* $\boldsymbol{B} = (B^1, B^2, \dots B^N)$ *is a* base-weighted true welfare index *and a set of numbers* $\boldsymbol{C} = (C^1, C^2, \dots C^N)$ *is a* current-weighted true welfare index *if:*

$$\begin{aligned} Q_{ij}^{CL} &\leq B^i/B^j \leq Q_{ij}^{L} \\ Q_{ij}^{P} &\leq C^i/C^j \leq Q_{ij}^{CU}, \quad \forall i,j = 1,\dots,N. \end{aligned}$$

From the above definition it is apparent that even if preferences are not common, then there may exist sets of numbers which satisfy the above inequalities. However, only if preferences are common can these numbers be interpreted as true welfare indexes. Having shown that a given set of demand data is consistent with the existence of a representative consumer, the next step is to construct the welfare indexes. By nature these are unobservable, but by the above definition, given that preferences are common then *any* set of numbers $\boldsymbol{B}$ and $\boldsymbol{C}$ which (by definition) are contained within the fixed-weight and classical bounds to the welfare indexes themselves qualify as true welfare indexes. In Chapter 2, three methods for constructing the numbers $\boldsymbol{B}$ and $\boldsymbol{C}$ using RP methods are reviewed.

It should be emphasised that constructing $\boldsymbol{B}$ and $\boldsymbol{C}$ using RP methods is equivalent to tightening the bounds to the Allen welfare indexes, since by the above definition, improved bounds to these welfare indexes are themselves true welfare indexes. It should be further noted that when $N = 2$ and preferences are common, the fixed-weight and classical bounds to the welfare indexes are true indexes. However, with $N > 2$ it is not necessarily the case that the fixed-weight and classical bounds will be true indexes.

A final remark relates to the informational content of the true indexes $\boldsymbol{B}$ and $\boldsymbol{C}$. It is possible to construct true indexes which do not have informational content in that they do not provide improved bounds to Q_{ij}^{LA} and Q_{ij}^{PA}. However the three RP methods for constructing true indexes reviewed in Chapter 2 in fact produce numbers with informational content in that they provide tight bounds to Q_{ij}^{LA} and Q_{ij}^{PA}.

Classical and fixed-weight bounds - Stone-Geary example

Above, the Allen welfare indexes for the two-good two-country Stone-Geary example were calculated. Under the more realistic situation where we do not know the form of the utility function, we can use the previous results to calculate fixed-weight and classical bounds to the welfare indexes for the example data:

$$\begin{array}{rclclcl} 0.4 & = & Q_{12}^{CL} & \leq & Q_{12}^{LA} & \leq & Q_{12}^{L} & = & 0.5874 \\ 0.5378 & = & Q_{12}^{P} & \leq & Q_{12}^{PA} & \leq & Q_{12}^{CU} & = & 0.8. \end{array}$$

Since the data for this example have been generated so as to conform with the Stone-Geary utility function, they must be consistent with common preferences. Hence, using Definition 1.1, the classical and fixed-weight bounds are themselves true welfare indexes.[24]

Example of non-common preferences

To finish this sub-section on true welfare indexes, the following is an example of two countries which do not share common preferences. Assume the demand data are $\mathbf{p}^1 = \{1,1\}$, $\mathbf{q}^1 = \{3,1\}$ and $\mathbf{p}^2 = \{1/2,1\}$, $\mathbf{q}^2 = \{1,2.5\}$. The classical and fixed-weight bounds for these data can be calculated as:

$$\begin{array}{rclclclclclcl} 0.667 & = & 2/3 & = & Q_{12}^{CL} & \leq & Q_{12}^{LA} & \leq & Q_{12}^{L} & = & 2.5/3 & = & 0.833 \\ 1.143 & = & 4/3.5 & = & Q_{12}^{P} & \leq & Q_{12}^{PA} & \leq & Q_{12}^{CU} & = & 4/3 & = & 1.333. \end{array}$$

However, since these data do not satisfy the condition in Proposition 1.1 ($Q_{12}^{L} < 1$ but $Q_{12}^{P} > 1$), then the true indexes do not exist. Thus, though it is possible to find numbers that satisfy the above inequalities, these numbers cannot be interpreted as true indexes since preferences are not common. Similarly, the classical and fixed-weight bounds themselves do not form true indexes (and in fact they really should not even be called bounds, since they do not bound anything).

[24] Note that since the budget sets for this particular example do not intersect, then the test of WARP has no power, and it is impossible to reject the null hypothesis that preferences are common whatever $\mathbf{q}^1$ and $\mathbf{q}^2$ might be observed (this issue is discussed in depth in Chapter 2).

1.4.4 A unique welfare index

From the above discussion, it is apparent that for a given data set that is consistent with common preferences, attempts to construct welfare comparisons between countries will be affected by issues relating to both *indeterminateness* and *non-uniqueness*. The indeterminateness arises because the utility function is unobservable, and hence only bounds to the welfare indexes can be constructed. However, a welfare comparison may also be affected by non-uniqueness in that the money metric utility ratios will depend on the choice of the reference price vector.

In certain contexts it may be appropriate to choose a particular reference price vector for use in the welfare comparison. An example might be the construction of a series of real income for a particular country over time, where prices are set to the first year in the data set. However, in the context of cross-country welfare comparisons, non-uniqueness of welfare comparisons is a considerable drawback since it may not be feasible or appropriate to use the price vector of a particular base country in the comparisons. One would have difficulty promoting the validity of country welfare rankings if the rankings constructed using U.S. prices, for example, are significantly different to those constructed using the prices of India.

However, this non-uniqueness does not exist when preferences are homothetic, and for this reason, considerable research in index number theory has focused on the case of homotheticity.

Proposition 1.2. EXISTENCE OF A UNIQUE WELFARE INDEX - VARIOUS: *There exists a unique index comparing the welfare of country i and j and defined by:* $Q_{ij}^{LA} = Q_{ij}^{PA} = U^i/U^j$, *if and only if preferences are homothetic.*

PROOF: We first show the sufficiency of homothetic preferences for the existence of a true welfare index. If the utility function is linearly homogenous, then it was shown above that $e(U, \mathbf{p}) = Ub(\mathbf{p})$, for some linearly homogeneous function $b(\mathbf{p})$. Hence, it follows that there exists a unique (i.e. price-independent) true welfare index which is the ratio of (money metric) utility levels, and the Laspeyres-Allen and Paasche-Allen quantity indexes are equal to this index.

To show necessity of homothetic preferences for the existence of a unique welfare index, we need to show that if there exists a unique index, then it must be that preferences are homothetic. The proof of this is fairly involved and is omitted here (see Samuelson and Swamy (1974, p.570) for an example of this proof).□

The existence of the unique welfare index when preferences are homothetic is shown graphically in Figure 1.4 for a two-good two-country example (where it is assumed that $p_2 = 1$ so that expenditures can be measured along the vertical axis). The expansion path for country i is $EP(\mathbf{p}^i)$ - with homotheticity the expansion path is a straight line through the origin. It is apparent from this figure that $Q_{21}^{LA} = 0B/0A$ and $Q_{21}^{PA} = 0D/0C$. Furthermore, the property of homotheticity allows us to show (using similar triangles) that $0B/0A = 0D/0C$ and hence $Q_{21}^{LA} = Q_{21}^{PA} = U^2/U^1$. Finally, it is apparent that $Q_{21}^{P} = 0D/0E < U^2/U^1$, that is, the Paasche index is the lower bound to the unique welfare index (and it can similarly be shown that the Laspeyres index is the upper bound to this index).

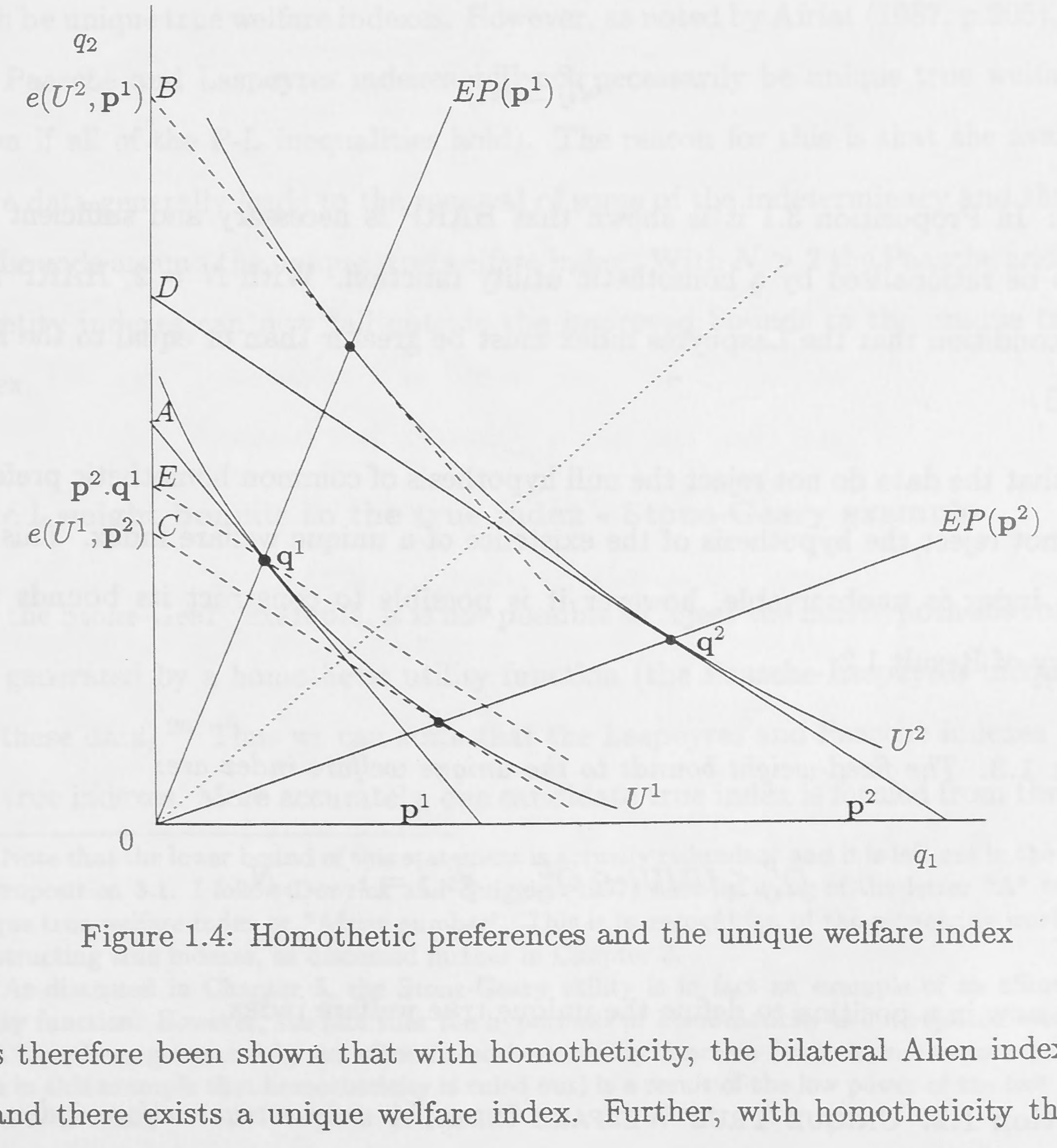

Figure 1.4: Homothetic preferences and the unique welfare index

It has therefore been shown that with homotheticity, the bilateral Allen indexes coincide and there exists a unique welfare index. Further, with homotheticity the Konüs and Malmquist welfare indexes also reduce to the ratio of utility levels U^i/U^j (since $e(U^i, \mathbf{p}^r) = U^i b(\mathbf{p}^r)$ and $d(\mathbf{q}^i, U^r) = u(\mathbf{q}^i)/U^r$) and thus are indifferent to the reference points used in their calculation.

The next question is naturally: how do we test a finite body of demand data for consistency with homothetic preferences? In Chapter 3 (Proposition 3.1) it is shown that a given set of data is consistent with the existence of a representative consumer who maximises a homothetic utility function if the data satisfy the Homothetic Axiom of Revealed Preference (HARP). For the case of $N = 2$ the test of common homothetic preferences reduces to the following condition:

Proposition 1.3. REVEALED PREFERENCE TEST OF COMMON HOMOTHETIC PREFERENCES WITH $N = 2$ - VARIOUS: *With $N = 2$, we do not reject the hypothesis that preferences across country i and country j are common and homothetic if and only if the data satisfy the Paasche-Laspeyres (P-L) inequality, that is:*

$$Q_{ij}^P \leq Q_{ij}^L.$$

PROOF: In Proposition 3.1 it is shown that HARP is necessary and sufficient for the data to be rationalised by a homothetic utility function. With $N = 2$, HARP reduces to the condition that the Laspeyres index must be greater than or equal to the Paasche index.□

Given that the data do not reject the null hypothesis of common homothetic preferences, we cannot reject the hypothesis of the existence of a unique welfare index. This unique welfare index is unobservable, however it is possible to construct its bounds using a corollary of Result 1.2:

Result 1.3. *The fixed-weight bounds to the unique welfare index are:*

$$Q_{ij}^P \leq U^i/U^j \leq Q_{ij}^L, \quad \forall i, j = 1, \ldots, N.$$

We are now in a position to define the unique true welfare index.

Definition 1.2. UNIQUE TRUE WELFARE INDEX *A unique true welfare index is a set* *of numbers* $\boldsymbol{A} = (A^1, A^2, \ldots A^N)$ *such that:*[25]

$$Q^P_{ij} \leq A^i/A^j \leq Q^L_{ij} \quad \forall i,j = 1,\ldots,N.$$

In Proposition 3.1 it is shown that the existence of the numbers $\boldsymbol{A}$ as defined above is necessary and sufficient for the data to be rationalised by a homothetic utility function. However, it should be noted that the inherent unobservability of the utility function and the resultant indeterminateness of the unique welfare index means that there will in fact be an infinite set of candidate unique true welfare indexes that are consistent with the data. As shown in Chapter 3, it is possible to use RP methods to calculate particular unique true welfare indexes which have informational content in that they form tight bounds to all possible unique true welfare indexes (and thus the range of indeterminacy to the unique unobservable welfare index U^i/U^j is reduced via the use of RP methods). When $N = 2$ and the P-L inequality holds, then the Paasche and Laspeyres indexes will both be unique true welfare indexes. However, as noted by Afriat (1987, p.205), for $N > 2$ the Paasche and Laspeyres indexes will not necessarily be unique true welfare indexes (even if all of the P-L inequalities hold). The reason for this is that the availability of more data generally leads to the removal of some of the indeterminacy and thus narrows the bounds around the unique true welfare index. With $N > 2$ the Paasche and Laspeyres quantity indexes can now fall outside the improved bounds to the unique true welfare index.

Fixed-weight bounds to the true index - Stone-Geary example

For the Stone-Geary example, it is not possible to reject the null hypothesis that the data are generated by a homothetic utility function (the Paasche-Laspeyres inequality holds for these data).[26] Thus we can state that the Laspeyres and Paasche indexes themselves are true indexes. More accurately, one candidate true index is formed from the Laspeyres

[25] Note that the lower bound of this statement is actually redundant and it is left out in the presentation in Proposition 3.1. I follow Dowrick and Quiggin (1997) here by using of the letter "A" to represent a unique true welfare index or "Afriat number". This is in recognition of the pioneering work of Afriat in constructing true indexes, as discussed further in Chapter 3.

[26] As discussed in Chapter 5, the Stone-Geary utility is in fact an example of an affine homothetic utility function. However, the fact that the hypothesis of homotheticity is not rejected even though the data have been generated by an affine homothetic utility function (and the subsistence parameters are such in this example that homotheticity is ruled out) is a result of the low power of the test which occurs when N is small.

index as $A_1 = \{A_1^1, A_1^2\} = \{0.587, 1\}$ and another candidate true index is formed from the Paasche index as $A_2 = \{A_2^1, A_2^2\} = \{0.538, 1\}$.

Example of data which do not satisfy common homothetic preferences

The data introduced above which do not satisfy common preferences ($\mathbf{p}^1 = \{1, 1\}$, $\mathbf{q}^1 = \{3, 1\}$ and $\mathbf{p}^2 = \{1/2, 1\}$, $\mathbf{q}^2 = \{1, 2.5\}$) will obviously also not satisfy common homothetic preferences. This can be seen by the fact that the Paasche index exceeds the Laspeyres index.

1.5 Multilateral Indexes

As noted by Samuelson and Swamy (1974, p.568), among others, the fundamental point about a welfare index is that itself must be a cardinal indicator of ordinal utility. The authors note that if a particular index were to satisfy all the axioms devised by Fisher (1922) except that of circularity, the index would fail as a cardinal indicator of utility. They illustrate this by showing that if a particular index Q_{ij} does not satisfy the circularity property, then the following could be possible:

$$Q_{20} = Q_{10} = 1$$

$$Q_{30} = Q_{10} = 1$$

$$Q_{32} > 1.$$

However, this would lead to the following intransitivity:

$$u(\mathbf{q}^2) = u(\mathbf{q}^1)$$

$$u(\mathbf{q}^3) = u(\mathbf{q}^1)$$

$$u(\mathbf{q}^3) > u(\mathbf{q}^2).$$

The moral of the story is summarised by Samuelson and Swamy (1974, p.576): "So long as we stick to the economic theory of index numbers, the circular test is required as is the property of transitivity itself. And this is regardless of homotheticity or nonhomotheticity." The central importance of circularity to the usefulness of a welfare index number in multilateral comparisons leads to the following definitions.

Definition 1.3. MULTILATERAL INDEX *A multilateral index is an index which satisfies the property of circularity.*

Definition 1.4. BILATERAL INDEX *A bilateral index is an index which is not a multilateral index.*

As noted above, the Laspeyres, Paasche and Ideal Fisher indexes do not satisfy circularity - they are all bilateral indexes.[27] However, there are three multilateral indexes proposed in the literature which are now reviewed.

[27] The failure of the Fisher index to satisfy the circularity test led Fisher to declare "..., therefore, a *perfect* fulfillment of this so-called circular test should really be taken as proof that the formula which fulfills it is erroneous." (1922, p.271)

1.5.1 Multilateral true indexes

The major focus of this thesis is on the construction of multilateral true indexes, which are now defined:

Definition 1.5. MULTILATERAL TRUE INDEX *A multilateral true index is a true index which also has the property of being a multilateral index.*

The above definition appears trivial, however it leads to an important result. With $N = 2$ it was shown above that if the P-L inequality is satisfied, then the Paasche and Laspeyres indexes are both unique true indexes. However, since these indexes do not satisfy circularity (which reduces to base-country invariance when $N = 2$), they are not multilateral true indexes, even when $N = 2$. In fact, it is apparent (and inherently reasonable) that a multilateral true index can only be constructed with $N > 2$.

It should be further noted that any unique true index constructed for $N > 2$ must satisfy circularity (this is apparent from Definition 1.2) and hence the condition for the existence of a multilateral true index is the same as the condition for the existence of a unique true index, namely the data must be consistent with the existence of a homothetic utility function. The construction of multilateral true indexes is described in detail in Chapter 3.

Having defined a multilateral true index, it remains to define a bilateral true index.

Definition 1.6. BILATERAL TRUE INDEX *A bilateral true index is a true index which is not a multilateral true index.*

It is apparent that with $N = 2$, if the P-L inequality is satisfied then the Paasche and Laspeyres indexes are both bilateral true indexes. Further, any set of numbers $\boldsymbol{B}$ or $\boldsymbol{C}$ defined in Definition 1.1 is a bilateral true welfare index. The construction of bilateral true welfare indexes is reviewed in Chapter 2.

1.5.2 Superlative indexes

While the multilateral true index defined above is the main focus of this thesis, it is important to review some of the "competing" multilateral indexes which are used in international comparisons. In this sub-section, a review of superlative index numbers is presented. This approach is due to Erwin Diewert and provides a method for measuring the cost-of-living and welfare which allows for substitution, yet does not require the estimation of a consumer demand system.[28]

While the Ideal Fisher index has certain desirable properties, the fact that it does not satisfy circularity makes it of limited use for cross-country comparisons. A multilateral version of the Ideal Fisher index is the EKS index of Eltetö and Köves (1964) and Szulc (1964):[29]

$$Q_{ij}^{EKS} = \prod_{k=1}^{N} \left(\frac{Q_{ik}^F}{Q_{jk}^F} \right)^{(1/N)} . \tag{1.4}$$

The EKS index is thus the geometric mean of the ratios of all N bilateral Fisher indexes, taking each country in turn as base.

Caves, Christensen, and Diewert (1982) have proposed an alternative multilateral index, which is similar in resemblance to the EKS index, but possesses superior theoretical properties. The basic "building block" for the CCD index is the bilateral Törnqvist index:

$$Q_{ij}^T = \prod_{l=1}^{K} \left(\frac{q_l^i}{q_l^j} \right)^{\frac{\omega_l^i + \omega_l^j}{2}}$$

The CCD index then multilateralises the Törnqvist index in the same way as the EKS index extends the Fisher index:

$$Q_{ij}^{CCD} = \prod_{k=1}^{N} \left(\frac{Q_{ik}^T}{Q_{jk}^T} \right)^{(1/N)} .$$

While the Ideal Fisher and Törnqvist indexes do not satisfy circularity, the EKS and CCD indexes do. Further, the EKS index exhibits a high degree of characteristicity by construction since it is the solution to the problem of finding an index number which satisfies circularity and minimises the sum of squared deviations from the Fisher indexes (Drechsler 1973). However both the EKS and CCD indexes fail to satisfy matrix consistency.

[28]See Diewert (1976, 1978, 1981, 1983) for further reading on this approach.

[29]This index was in fact first proposed by Gini (1931). The EKS index is used by the OECD and by Eurostat (the Statistical Office of the European Union) to produce purchasing-power-parity-corrected real income data for their member countries.

Definition 1.7. EXACT QUANTIY INDEX[30] *A quantity index Q_{ij} is exact for a given utility function $u(\mathbf{q})$ if $Q_{ij} = u(\mathbf{q}^i)/u(\mathbf{q}^j)$, where $\mathbf{q}^i$ is the Marshallian demand for $u(\mathbf{q})$.*

In other words, a quantity index number is exact if it exactly equals the ratio of the utility functions whenever the data is consistent with microeconomic maximising behaviour. The EKS index is exact when the utility function is a homogeneous quadratic, while the CCD index is exact for a homogeneous translog utility function.[31]

Definition 1.8. SUPERLATIVE INDEX *A superlative index is an index which is exact for a utility function which is flexible (in that it provides a second-order approximation to an arbitrary twice-differentiable utility function).*

The popularity of the EKS and CCD indexes in empirical work can therefore be attributed to the fact that they are both superlative index numbers. However, despite their popularity, the use of superlative indexes in multilateral comparisons can be questioned on three grounds. First, while these indexes are exact for particular (flexible) specifications of the utility function, there is no testing as to whether the data are consistent with these utility functions.[32] In contrast, multilateral true indexes constructed under the RP approach only exist when the data has been found to be consistent with homothetic preferences. Despite being exact for homogeneous utility functions (which by definition are homothetic), neither the EKS nor CCD indexes are necessarily multilateral true indexes (as defined above) in that they do not always lie within the fixed-weight bounds to the true index. If the data were exactly consistent with these specific underlying (homothetic) utility functions, the EKS/CCD indexes would be multilateral true indexes.

Second, while the EKS and CCD indexes exhibit circularity, Neary (2000) has noted that this is only a statistical property (it is an artifact of the averaging process used in their construction), and does not reflect an underlying preference structure which necessarily leads to circularity.[33] For example, the EKS index is a set of real numbers

[30]Note that the property of exactness was originally proposed for any aggregator function (that is, it applies to both utility functions and production functions).

[31]In the context of superlative indexes, homogeneity is taken to mean *linear* homogeneity.

[32]Afriat originally made this criticism of the finding by Byushgens (1925) that the Fisher index is exact for homogeneous quadratic preferences.

[33]In practice it is found that Fisher indexes rarely exhibit circularity, suggesting that the underlying preferences are *not* homogeneous quadratic. Thus, imposing a specific homothetic functional form is likely to be too restrictive.

which are constructed as averages of N ratios of Fisher indexes and the EKS only exhibits circularity because this is an inherent property of the real number system. In contrast, the multilateral true index introduced above exhibits circularity by virtue of the fact that it is constructed when preferences are shown to be homothetic.

Finally, Neary (2000) has argued that the property of exactness in fact makes the EKS and CCD indexes redundant and thus unsuitable for use in multilateral comparisons. The Neary Critique of the use of the EKS and CCD indexes in multilateral comparisons is now summarised.

Result 1.4. Exactness of the Fisher index - Konüs and Byushgens (1926): *The Fisher index is exact when the utility function is a homogeneous quadratic:* $U = (\mathbf{q}'\mathbf{A}\mathbf{q})^{1/2}$, $\mathbf{A}$ *symmetric.*

Since the homogeneous quadratic utility function is homothetic, the fact that the Fisher index is exact means that it is equal to the utility ratios:

$$Q_{ij}^{F} = U^i/U^j. \tag{1.5}$$

As noted by Diewert (1976), the quadratic utility function is flexible in that it provides a second-order approximation to an arbitrary twice-differentiable linearly homogeneous utility function. Hence, the Fisher index is superlative. The fact that the EKS index is the multilateral version of the superlative bilateral Fisher index would appear to justify its use in multilateral comparisons. However, the following result does not support this.

Result 1.5. Exactness of the EKS index - Neary (2000): *The EKS index is exact when the utility function is a homogeneous quadratic.*

This result follows from the substitution of (1.5) into (1.4):

$$Q_{ij}^{EKS} = \prod_{k=1}^{N} \left(\frac{U^i/U^k}{U^j/U^k} \right)^{(1/N)} = U^i/U^j.$$

The above shows that the EKS index is superlative in that when preferences are homogeneous quadratic it is equal to the ratio of utility levels, just as the Fisher index is. However, Neary (2000) argues that it is precisely this equality that provides an argument against the use of the EKS index for multilateral comparisons. When preferences are homogeneous quadratic, the EKS index is redundant since it is equal to the bilateral Fisher index. In practice, we generally find an inequality between the Fisher and EKS indexes (and, relatedly, that the Fisher index does not satisfy circularity), thus indicating that the demand data could not be generated by homogeneous quadratic preferences. As mentioned above, the EKS index will by construction exhibit circularity, unlike the Fisher index. However, this is a statistical property and does not imply that the underlying preference structure has led to circularity.

The Neary Critique of the CCD index is similar to the critique of the EKS index.

Result 1.6. EXACTNESS OF THE TÖRNQVIST INDEX - DIEWERT (1976): *The Törnqvist index is exact when the utility function is a homogeneous translog:* $\ln U = a_0 + \mathbf{a}' \ln \mathbf{q} + \frac{1}{2}(\ln \mathbf{q})' \mathbf{A} \ln \mathbf{q}$.

The translog is also flexible functional form, and is in fact more general than the quadratic; this suggests that the Törnqvist bilateral is even more "superlative" than the bilateral Fisher index. For this reason, the CCD index, which is the multilateral version of the Törnqvist bilateral index, has been advocated in multilateral comparisons. However, Neary (2000) again throws doubt on the use of the CCD index.

Result 1.7. EXACTNESS OF THE TÖRNQVIST INDEX - NEARY (2000): *The CCD index is exact where the utility function is homogeneous translog.*

The above arguments cast serious doubt on the use of the EKS and CCD indexes in international comparisons research. However, these indexes are (and probably will continue to be) popular in cross-country comparative research.[34]

1.5.3 The Geary index

The Geary index, which is used by the ICP to compute purchasing-power-parity-corrected exchange rates and real income measures, is constructed in an entirely different manner to the construction of the other multilateral indexes discussed so far.[35] The index was

[34]One extension of the use of the CCD index is the stochastic approach of Selvanathan and Rao (1994). This approach involves the use of an averaging process to eliminate random errors which make each price relative deviate from the overall price index. The method gives an estimate of the underlying signal, the price index, and provides standard errors for index numbers.

[35]Khamis(1970, 1972) derived conditions for the Geary method to give a unique, positive solution and advocated the use of index numbers based on the Geary method. The Geary index is also known as the Geary-Khamis index, however Neary (1996) has argued that this may be over-generous to Khamis. The Geary index is used extensively in applied work and was popularised by its use in the Penn World Tables. The presentation here is based on Neary (2000).

first suggested by Geary (1958), and is constructed under the assumed existence of "world prices" $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)$ and "true" exchange rates $\boldsymbol{\epsilon} = (\epsilon^1, \ldots, \epsilon^N)$. The true exchange rates are Laspeyres price indexes, which compare the world prices with the prices of each country in turn:[36]

$$\epsilon^i = \frac{\sum_l \pi_l q_l^i}{\sum_l p_l^i q_l^i}, \quad i = 1, \ldots, N. \tag{1.6}$$

It is therefore apparent that each country's real income is the same whether valued at world prices ($\sum_l \pi_l q_l^i$) or valued at domestic prices, converted at the true exchange rates ($\epsilon^i \sum_l p_l^i q_l^i$). The world prices are defined by the requirement that total world spending on commodity l is the same whether valued at its world price ($\pi_l \sum_i q_l^i$) or at domestic prices converted at the true exchange rates ($\sum_i \epsilon^i p_l^i q_l^i$):

$$\pi_l = \frac{\sum_i \epsilon^i p_l^i q_l^i}{\sum_i q_l^i}, \quad l = 1, \ldots, K. \tag{1.7}$$

Solving simultaneously for $\boldsymbol{\epsilon}$ and $\boldsymbol{\pi}$, one can then calculate the real income of each country at world prices:

$$x_G^i = \epsilon^i x^i = \sum_l \pi_l q_l^i, \quad i = 1, \ldots, N.$$

The Geary real income measure implies a set of indexes, $Q_{ij}^G = x_G^i / x_G^j$.

Thus, with the Geary index, the welfare of each country is measured at world prices, $\boldsymbol{\pi}$. The Geary index exhibits circularity (and hence is a multilateral index), but only because the real income of each country is measured at the same consistent set of prices, $\boldsymbol{\pi}$.[37] However the world prices are artificial in that they are not faced by anyone (by construction, they reflect an "average international" price structure).

One of the advantages of the Geary index is that it exhibits matrix consistency. However, the Geary index is a fixed-weight index and thus does not allow for any substitution

[36]Following the U.S. convention, the ICP defines the true exchange rates or "purchasing power parities" as the inverse of the definition presented above; Neary (2000) follows the U.K. convention since it facilitates the matrix algebra.

[37]As noted by Samuelson and Swamy (1974, p.577) in the absence of homotheticity, circularity is still satisfied *in a sense* if the welfare comparisons are made using a particular price vector. For example, if $\mathbf{p}^r$ is used in the welfare comparison, then circularity will hold:

$$Q_{ij}^{A,r} = \frac{e(U^i, \mathbf{p}^r)}{e(U^j, \mathbf{p}^r)} = \frac{e(U^i, \mathbf{p}^r)}{e(U^k, \mathbf{p}^r)} \frac{e(U^k, \mathbf{p}^r)}{e(U^j, \mathbf{p}^r)} = Q_{ik}^{A,r} Q_{kj}^{A,r}.$$

However, there is no reason to choose $\mathbf{p}^r$ over any other possible price vector for the welfare comparison; homotheticity is required for "true" circularity to hold.

in consumption. The Geary reference price index π tends to be similar to the price vectors of richer countries; as discussed in Section 1.3.2, the presence of the Gerschenkron effect leads to the Geary index underestimating per capita income differences across countries. Using 1985 ICP data on 64 countries, Hill (2000) found that the Geary index overestimates per capita income for some poorer countries, relative to richer countries, by as much as 70 percent.

The Geary method and measures of inequality and convergence

Relatedly, the level of cross-country inequality is also underestimated when the Geary per capita income measure is used. Dowrick and Quiggin (1997) show that the variance of log GDP from the Penn World Table (PWT), which uses the Geary aggregation method, is relatively low (compared with measures of variance constructed using poor country price vectors), reflecting the fact that the Geary international price vector resembles the prices of richer countries.[38] Hill (2000) found for the 1985 ICP data that the Gini coefficient for real per capita income constructed using the Geary index was 0.57, which is toward the lower end of the range of Ginis calculated using all of the different country price vectors. Since the Geary international price vector is closer to the price vector of richer countries, it follows from the discussion in Section 1.3.3 that σ convergence calculated using GDP constructed with the Geary method will be underestimated if prices are also converging. Dowrick and Quiggin (1997) calculate that the variance of log (PWT) GDP falls by 0.016 between 1980 and 1990; this is substantially less than the fall of 0.023 which is found using their preferred measure of country welfare, the Ideal Afriat Index (which is discussed in detail in Chapter 3).

Geary-Allen International Accounts System

From the above, it appears that the fact that the Geary index exhibits substitution bias seriously compromises its use in international comparisons research. However, recent important research by Neary (2000) aims to redress this drawback, while preserving

[38]Dowrick and Quiggin (1997) also present measures of inequality and convergence calculated using GDP constructed by the ICP. However, the ICP used the Geary method in 1980 and shifted to the EKS method in 1990, the change in aggregation methods makes comparisons between years more difficult. For this reason, the focus here is on inequality and convergence using the PWT GDP series.

the spirit of the Geary method. Neary (2000) has proposed a new index, the Geary-Allen International Accounts (GAIA) System, which uses estimated parameters from the Quadratic Almost Ideal Demand System (QUAIDS) of Banks, Blundell, and Lewbel (1997) to incorporate substitution in consumption into the Geary index.

A complete discussion of the GAIA is beyond the scope of this thesis; a brief outline follows.[39] The first step is to replace the fixed-weight Geary exchange rates (1.6) with their true equivalents, which are known as the Geary-Konüs exchange rates:

$$E^i = \frac{e(U^i, \mathbf{\Pi})}{e(U^i, \mathbf{p}^i)} = \frac{\sum_l \Pi_l q_l^{*i}}{\sum_l p_l^i q_l^i}, \quad i = 1, \dots, N, \tag{1.8}$$

where q_l^{*i} is the "virtual" or imputed quantity that the representative consumer would choose if he or she had country i's level of utility and faced the world prices $\mathbf{\Pi} = \{\Pi_1, \dots, \Pi_K\}$, and is calculated by applying Shephard's Lemma:[40]

$$q_l^{*i} = e_l(U^i, \mathbf{\Pi}),$$

where $e_l(U^i, \mathbf{\Pi}) = \partial e(U^i, \mathbf{\Pi})/\partial \Pi_l$. The second step is to require that the world prices satisfy the Geary aggregation conditions and (1.7) is therefore replaced with:

$$\Pi_l = \frac{\sum_i E^i p_l^i q_l^i}{\sum_i q_l^{i*}}, \quad l = 1, \dots, K. \tag{1.9}$$

Solving simultaneously for $\mathbf{\Pi}$ and $\boldsymbol{E} = (E^1, \dots, E^N)$ - using an approach to be described further below - one can then calculated the GAIA real income of each country:

$$x_{GA}^{*i} = E^i x^i = \sum_l \Pi_l q_l^{*i} = e(U^i, \mathbf{\Pi}), \quad i = 1, \dots, N.$$

The Geary-Allen true indexes of real income are:

$$Q_{ij}^{GA} = \frac{x_{GA}^{*i}}{x_{GA}^{*j}} = \frac{e(U^i, \mathbf{\Pi})}{e(U^j, \mathbf{\Pi})} \quad \forall i, j = 1, \dots, N.$$

Neary (2000) proves that there exists a solution to (1.8) and (1.9) with all E^i, Π_l and q_l^{i*} strictly positive. This solution is arrived at via a tâtonnement process which can be described as follows. First, the Geary system (1.6) and (1.7) is solved to calculate the Geary prices (see Appendix A for details of how this system may be solved). Next, the

[39] A difference with the presentation here, compared with that in Neary (2000), is that identical tastes worldwide are assumed from the outset; Neary (2000) introduces this assumption later.

[40] This footnote has been removed.

Geary prices are used to estimate a QUAIDS model - the estimate parameters of which are then used to calculate the virtual quantities q_l^{i*} corresponding to the Geary prices. These virtual quantities are then used in the GAIA system (1.8) and (1.9) which is solved to find an estimate of $\mathbf{\Pi}$.

This estimate of GAIA world prices is then used in a second round estimation of the QUAIDS model and the implied virtual quantities. This iterative process continues until there is no further change in the GAIA prices and the final estimate of these prices is used for the construction of the Geary-Allen index of real income. Neary (2000) proves that there exists a unique tâtonnement path from the Geary prices to the GAIA prices (however, it is not proved that this GAIA price vector is itself unique).

The GAIA System and measures of inequality and convergence

Using 1980 ICP data, Neary (2000) calculates Geary-Allen indexes using several different specifications of the QUAIDS model and compares these indexes with the EKS and Geary index. By construction, the Geary-Allen index does not exhibit substitution bias and consequently, the various Geary-Allen indexes constructed by Neary (2000) imply a less compressed distribution of world income, compared with the findings for the Geary and EKS index. Neary (2000) hypothesises that the use of the Geary-Allen index will similarly produce differing conclusions regarding convergence of incomes; it would be of interest to estimate the GAIA System for 1990 ICP data and see whether the implied σ convergence is closer to that found by Dowrick and Quiggin (1997) using their Ideal Afriat Index.

1.6 Conclusions

In this chapter, a review of the axiomatic and economic approaches to welfare measurement was presented. The RP approach to welfare measurement, which is an example of the economic approach to index number construction, was outlined. A new framework for constructing bilateral true welfare indexes was introduced, and this framework is used in Chapter 2 where three methods for constructing bilateral true welfare indexes are reviewed. The multilateral true welfare index, which is the main focus of this thesis, was also introduced and compared with other multilateral indexes which have been proposed in the literature. Further results relating to the existence and construction of the multilateral true index are presented in Chapter 3.

Chapter 2

The Revealed Preference Approach to Welfare Measurement

2.1 Introduction

Under the money metric concept of utility, if a particular bundle of goods consumed by country A costs more than the bundle consumed by country B, when costs are measured using the same prices, then country A has higher utility than country B. There is an implicit assumption in money metric welfare comparisons that preferences are common across the countries being compared.[1] If this were not the case, and different countries had different utility maps, then one could not be certain that a bundle costing more imparted higher utility.

In Chapter 1, money metric welfare indexes constructed under the economic approach to welfare measurement were introduced. It was shown that if preferences across observational units are common, then the Allen welfare indexes exist. These welfare indexes are not unique - the welfare between two countries can be compared using either the (base-weighted) Laspeyres-Allen or (current-weighted) Paasche-Allen welfare index. Further, while these indexes exist when preferences are common, since we do not have complete information on consumer preferences, they are not directly observable. However, it was shown in Chapter 1 that bounds to the Allen welfare indexes are provided by the classical and fixed-weight quantity indexes.

[1]There is also an assumption of monotonicity of preferences.

As discussed in Chapter 1, for a given data set that is consistent with common preferences, a true index can be constructed as a set of numbers the ratios of which are contained within the classical and fixed-weight bounds to the Allen welfare indexes. A true index thus defined will not satisfy the property of circularity, that is, it will not be a multilateral index. Hence such an index is known as a *bilateral true index.* If a given data set does not reject the hypothesis that preferences are common and homothetic, however, then there will exist a unique true welfare index, which also has the property of being a multilateral index. Multilateral true indexes are the focus of Chapter 3.

The revealed preference (RP) approach to welfare measurement employs revealed preference methods for constructing bilateral true welfare indexes.[2] RP methods can be used to construct bilateral true indexes with informational content in that they form tight bounds to all possible bilateral true indexes. There are three main characteristics of the RP approach: (1) It is explicitly based on the utility maximisation framework (and therefore is an example of the economic approach to index number construction); (2) common preferences across countries are *hypothesised and tested for*[3] and (3) the tests of common preferences are conducted using revealed preference relations.

In Section 2.2, RP tests of common preferences are reviewed. The RP approach to testing for common general preferences originated with the work of Afriat (1967), Diewert (1973) and Varian (1982), and the most easily implemented version of this test is that of the Generalised Axiom of Revealed Preferences (GARP). If a data set is found to satisfy GARP, then we cannot reject the hypothesis that preferences are common and hence there exists a bilateral true welfare index.

One of the problems with the GARP test of common preferences is that it tends to have low power (in that it is difficult to reject the null hypothesis of common preferences). A related consequence of the low power of GARP is that bilateral true indexes constructed

[2]The methods which come under the RP approach have been referred to as "nonparametric" methods by other authors since they generally do not involve an assumption about the form of preferences. However, this term is not used in this thesis for the following reason. One of the methods described in this chapter (that of Blundell, Browning, and Crawford (1998)) expansion path information to improve the power of the RP test of common preferences. While Blundell, Browning, and Crawford (1998) estimate nonparametric Engel curves to obtain this expansion path information, their method could equally be used with a parametric specification of the Engel curve (as is done in this chapter). In such a situation, the term "nonparametric" would be a misnomer.

[3]An assumption of common tastes is imposed in the construction of superlative indexes (EKS and CCD indexes), and also the Geary index. Note that under the RP approach intra-national preferences are assumed to be common and to satisfy aggregation conditions necessary for the existence of a representative consumer.

related consequence of the low power of GARP is that bilateral true indexes constructed under the RP approach do not tend to be much tighter than the bounds provided by the classical and fixed-weight quantity indexes. In Section 2.3 it is shown how the power of the test of GARP can be improved (and, relatedly, the bounds to bilateral true indexes which are constructed via this approach can be tightened) if information on expansion paths is used. In particular, recent work by Blundell, Browning, and Crawford (1998) on incorporating expansion path information into RP tests of general common preferences is reviewed.

Three methods for constructing the bounds to bilateral true indexes are reviewed in section 2.3 (see Figure 2.1); one of the innovations of this chapter is the presentation of a framework for comparing these three methods. The GARP bounds of Varian (1982) are constructed without any information about the expansion path, and consequently are not particularly tight. In contrast, the improved GARP bounds of Blundell, Browning, and Crawford (1998) incorporate estimated expansion paths and are significantly tighter than the GARP bounds. Afriat (1987) has shown that given the existence of a utility function (i.e., given the data share common preferences), then the family of utility functions consistent with the data may be bounded by the inner and outer envelope utility functions. Chavas and Cox (1997) have shown how the bilateral true indexes can be approximated using the Afriat envelope functions and one of the findings in this chapter is that even when no information about the expansion path is used, the bilateral true indexes constructed from the Afriat envelope functions are improvements on the GARP bounds of Varian (1982).

The structure of this chapter is as follows. In Section 2.2, RP tests of common preferences are presented and applied in an example. In Section 2.3, it is shown how the power of the RP test can be improved via the use of expansion path information. Three RP methods for constructing bilateral true welfare indexes are reviewed in Section 2.4. Section 2.5 concludes this chapter.

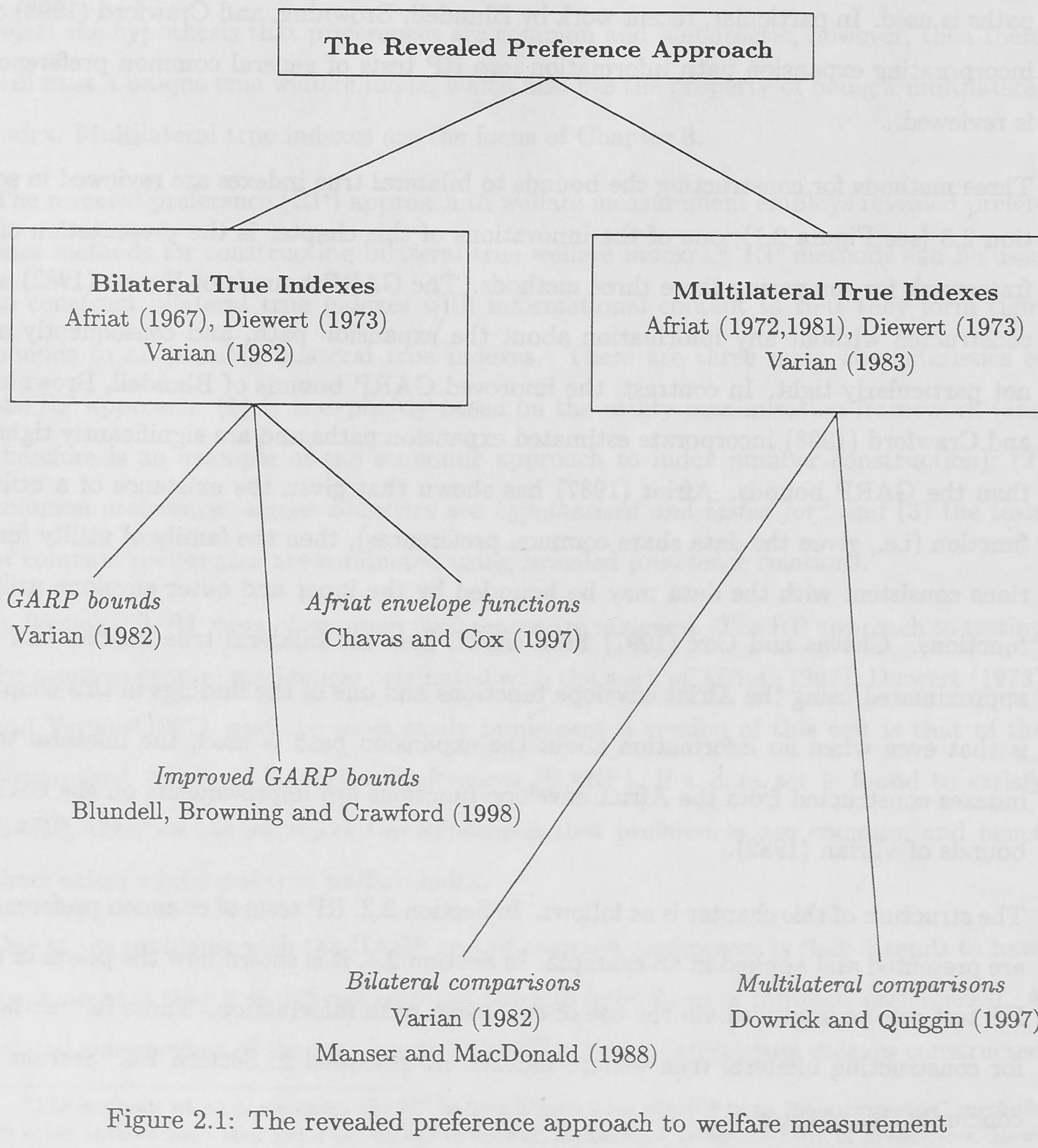

Figure 2.1: The revealed preference approach to welfare measurement

2.2 The RP Tests of Common General Preferences

In this section, the RP preference test of common preferences is reviewed.[4] This tests for the existence of a representative consumer who maximises general utility function. The three methods for conducting this test are first described and then implemented in an example.

2.2.1 Preliminaries

Revealed preference theory is associated with the work of Samuelson (1948), Houthakker (1950), Richter (1966), Afriat (1976) and others. Before describing the RP test of the representative consumer hypothesis, it is useful to introduce some general concepts related to revealed preference analysis.

The following definitions are used:

1. $\mathbf{q}^i$ is *directly revealed preferred* to $\mathbf{q}$ ($\mathbf{q}^i R^D \mathbf{q}$) if $\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}$;
2. $\mathbf{q}^i$ is *strictly directly revealed preferred* to $\mathbf{q}$ ($\mathbf{q}^i P^D \mathbf{q}$) if $\mathbf{p}^i.\mathbf{q}^i > \mathbf{p}^i.\mathbf{q}$;
3. $\mathbf{q}^i$ is *revealed preferred* to $\mathbf{q}$ ($\mathbf{q}^i R \mathbf{q}$) if $\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}^j, \mathbf{p}^j.\mathbf{q}^j \geq \mathbf{p}^j.\mathbf{q}^k, ..., \mathbf{p}^m.\mathbf{q}^m \geq \mathbf{p}^m.\mathbf{q}$ for some sequence of observations $(\mathbf{q}^i, \mathbf{q}^j, ..., \mathbf{q})$. In this case the relation R is said to be the *transitive closure* of the relation R^D;
4. $\mathbf{q}^i$ is *strictly revealed preferred* to $\mathbf{q}$ ($\mathbf{q}^i P \mathbf{q}$) if $\exists$ observations $\mathbf{q}^j$ and $\mathbf{q}^k$ such that $\mathbf{q}^i R \mathbf{q}^j, \mathbf{q}^j P^D \mathbf{q}^k, \mathbf{q}^k R \mathbf{q}$.

The simplest version of the RP test of common preferences involves the Weak Axiom of Revealed Preference (WARP), the definition of which follows (also see Figures 2.2 and 2.3).

Definition 2.1. *The data satisfy WARP if* $\mathbf{q}^i R^D \mathbf{q}^j$ *implies not* $\mathbf{q}^j R^D \mathbf{q}^i$*. In other words* $\mathbf{q}^i R^D \mathbf{q}^j$ *implies* $\mathbf{p}^j.\mathbf{q}^j < \mathbf{p}^j.\mathbf{q}^i$.

[4] Much of this summary is drawn from Varian (1982). A major simplification in the presentation here is that comparisons are made only between consumption bundles which have been observed; the presentation in Varian (1982) is more general as it allows for comparisons between observed and hypothetical consumption bundles. Also, the presentation here incorporates relevant results from Knoblauch (1992).

WARP therefore rules out inconsistencies of preferences. Thus if $\mathbf{q}^i$ was chosen in preference to $\mathbf{q}^j$ (that is, $\mathbf{q}^j$ was affordable at $\mathbf{p}^i$), then it must be that $\mathbf{q}^i$ was not affordable at $\mathbf{p}^j$ (i.e. $\mathbf{q}^j$ was not chosen in preference to $\mathbf{q}^i$).

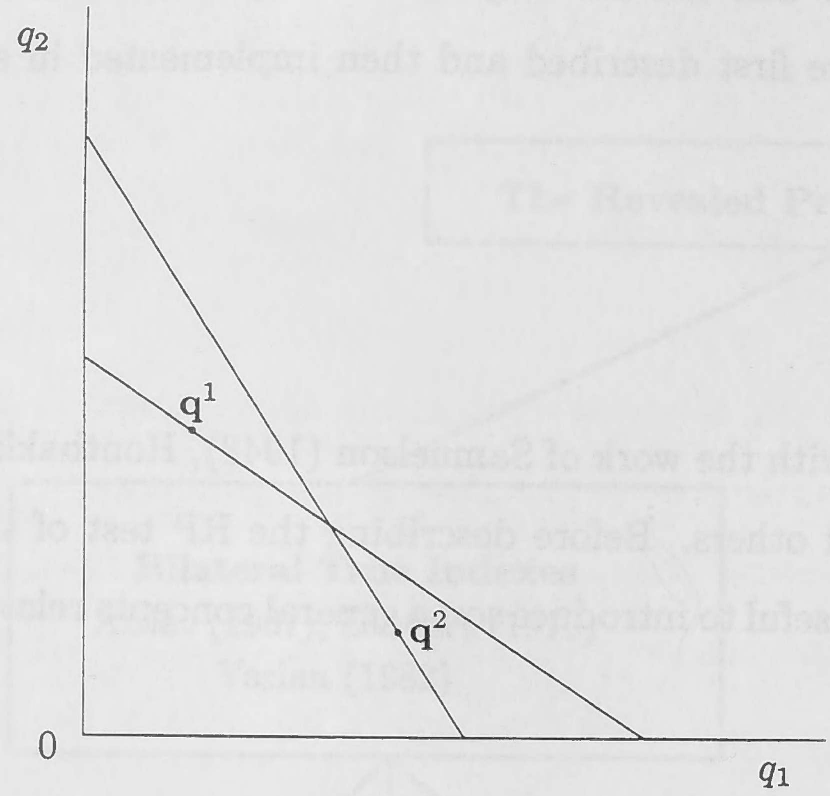

Figure 2.2: WARP is rejected

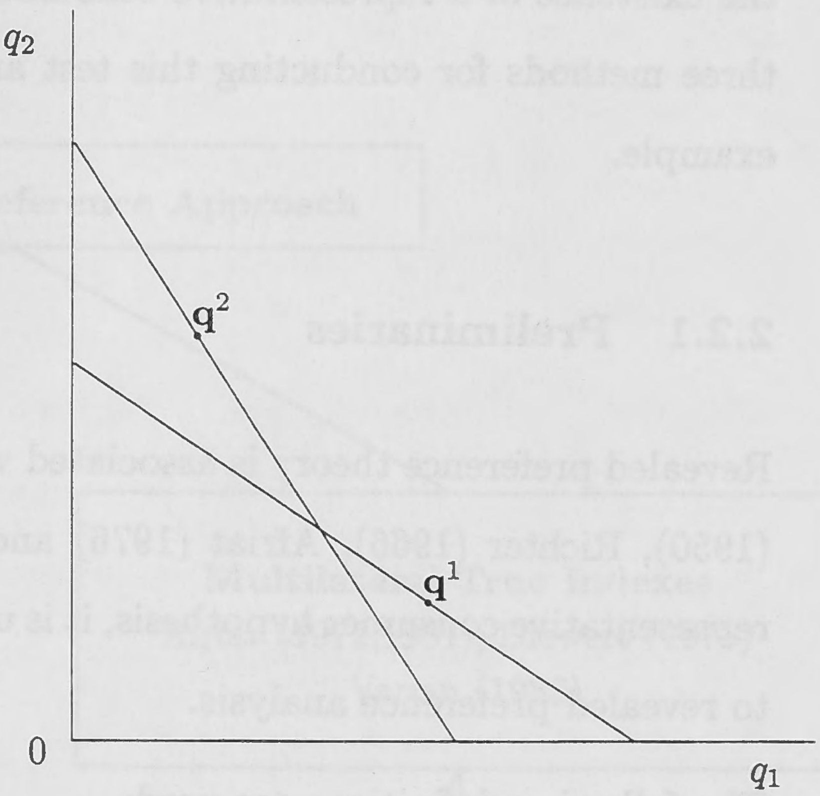

Figure 2.3: WARP is satisfied

The more general (and empirically relevant) version of the RP test of common preferences involves the Generalised Axiom of Revealed Preferences (GARP) the definition of which follows (and examples are in Figures 2.4 and 2.5).

Definition 2.2. *The data satisfy GARP if* $\mathbf{q}^i R \mathbf{q}^j$ *implies not* $\mathbf{q}^j P^D \mathbf{q}^i$. *In other words* $\mathbf{q}^i R \mathbf{q}^j$ *implies* $\mathbf{p}^j.\mathbf{q}^j \leq \mathbf{p}^j.\mathbf{q}^i$.[5]

Intuitively, GARP also rules out preference reversals. If $\mathbf{q}^i$ was chosen in preference to $\mathbf{q}^j$ at a particular set of prices, then if a new set of prices still allows both $\mathbf{q}^j$ and a bundle which is not inferior to $\mathbf{q}^i$ to be chosen, $\mathbf{q}^j$ will *not* be chosen.

The revealed preference relation R summarises all of the preference information contained in the demand data. Having defined a revealed preferred relation for a given set of data, the next step is to use this relation to associate utility functions with the data set via the concept known as rationalisation.

[5] There are other axioms of revealed preference, most notably the Strong Axiom of Revealed Preference (SARP). The data satisfy SARP if $\mathbf{q}^i R \mathbf{q}^j$ and $\mathbf{q}^i \neq \mathbf{q}^j$ implies not $\mathbf{q}^j R \mathbf{q}^i$. It can be shown that SARP implies GARP, but not vice-versa. SARP and WARP both require single-valued demand functions (i.e. the indifference curves cannot have any flat spots), while GARP is compatible with multi-valued demand functions, and for this reason it is more general.

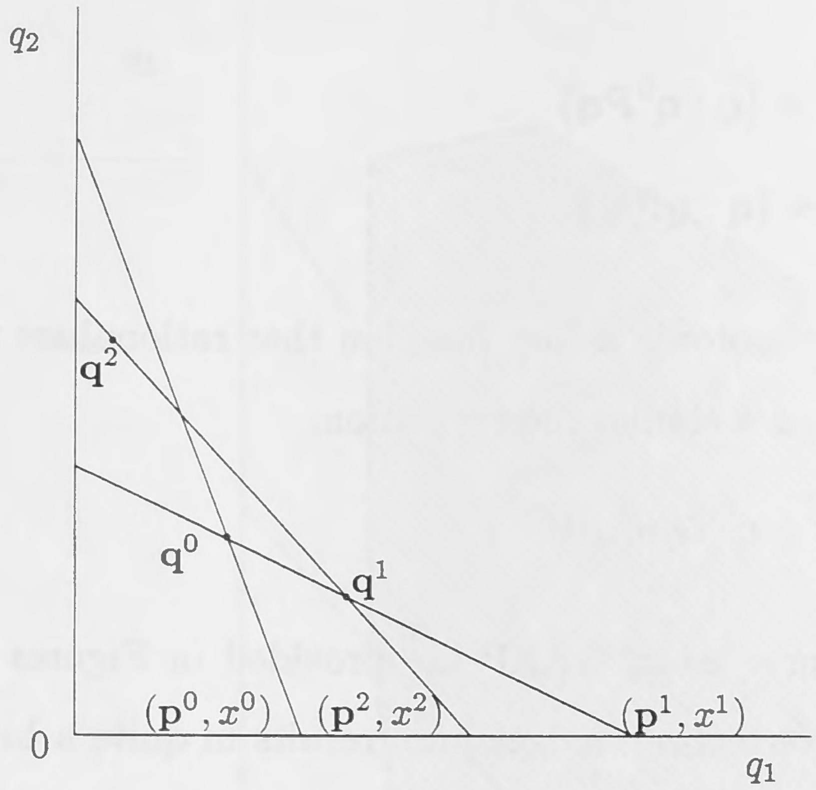

Figure 2.4: GARP is rejected

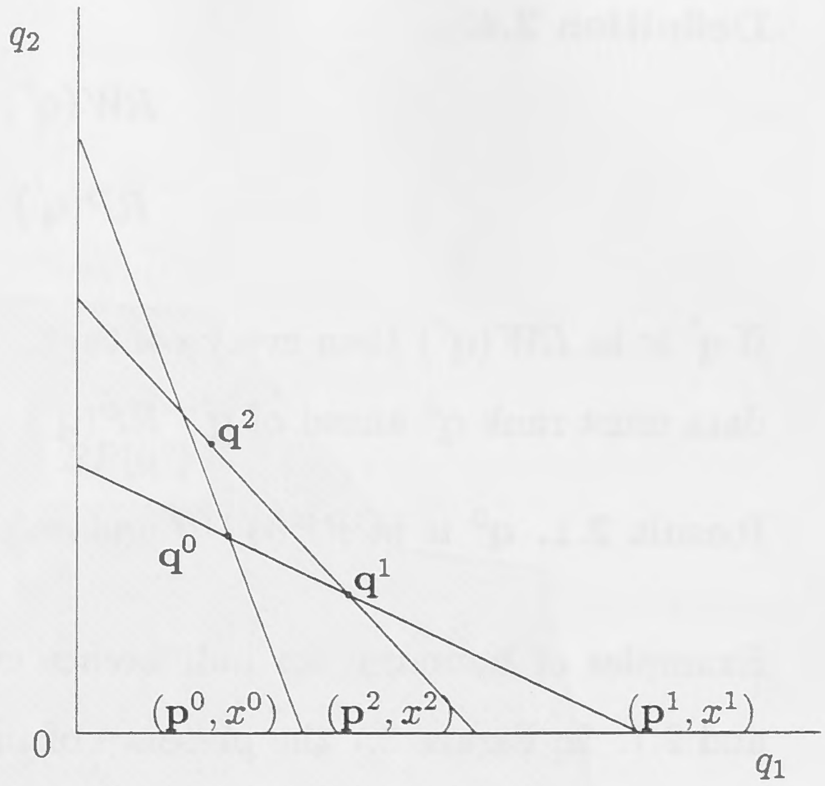

Figure 2.5: GARP is satisfied

Definition 2.3. *A utility function $w(\mathbf{q})$ rationalises the data $(\mathbf{p}, \mathbf{q})$ if $w(\mathbf{q}^i) \geq w(\mathbf{q})$ $\forall$ $\mathbf{q}$ such that $\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}$ i.e. $w(\mathbf{q})$ rationalises the observed behaviour if it achieves its maximum value on the budget set at the chosen bundles.*[6]

Broadly speaking, if a utility function $w(\mathbf{q})$ rationalises the data $(\mathbf{p}, \mathbf{q})$, then $w(\mathbf{q})$ could have generated the data. Note, however, that economists generally impose certain regularity conditions on the allowable utility functions that may rationalise a given set of demand data. For example, we may restrict $w(\mathbf{q})$ to be of a class of nonsatiated, monotonic and concave utility functions.[7] Afriat's theorem of the representative consumer hypothesis, which is presented below, implies that with no loss of generality such restrictions can always be imposed on the utility function. Conversely, with a finite amount of demand data that satisfy GARP, it is impossible to find violations of these restrictions.[8]

The RP approach to constructing bilateral bounds involves using the concepts introduced above to establish sets of observations which are revealed preferred and revealed worse than a given observation. The set of observations revealed worse to $\mathbf{q}^0$ and revealed preferred to $\mathbf{q}'$ are defined, respectively:

[6] The notation used here distinguishes an observable utility function $w(\mathbf{q})$ that rationalises the data and is one of an infinite number of functions that *could* have generated the data from the underlying unobservable utility function $u(\mathbf{q})$ that *did* generate the data.

[7] A utility function is *monotonic* if $\mathbf{q}^i \geq \mathbf{q}^j \Rightarrow w(\mathbf{q}^i) \geq w(\mathbf{q}^j)$ and is *locally nonsatiated* if arbitrarily near each $\mathbf{q}$ there is a $\mathbf{q}^j$ such that $w(\mathbf{q}^j) \geq w(\mathbf{q})$.

[8] The main point to note here is that the GARP test of common preferences does not rule out the existence of a utility function which does not have well behaved properties. However GARP being satisfied indicates that the data *could have been* generated by a well-behaved utility function.

Definition 2.4.

$$RW(\mathbf{q}^0) = \{\mathbf{q} : \mathbf{q}^0 P \mathbf{q}\}$$

$$RP(\mathbf{q}^{'}) = \{\mathbf{q} : \mathbf{q} P \mathbf{q}^{'}\}.$$

If $\mathbf{q}^{'}$ is in $RW(\mathbf{q}^0)$ then every concave, monotonic utility function that rationalises the data must rank $\mathbf{q}^0$ ahead of $\mathbf{q}^{'}$. $RP(\mathbf{q}^{'})$ has a similar interpretation.

Result 2.1. $\mathbf{q}^0$ *is in* $RP(\mathbf{q}^{'})$ *if and only if* $\mathbf{q}^{'}$ *is in* $RW(\mathbf{q}^0)$.

Examples of bounding an indifference curve using GARP are provided in Figures 2.6 and 2.7. In Figure 2.7 the presence of intersecting budget lines results in quite a bit of information about the indifference curve passing through $\mathbf{q}^0$; it can't intersect $RW(\mathbf{q}^0)$ or $RP(\mathbf{q}^0)$, and hence it must lie between the two.

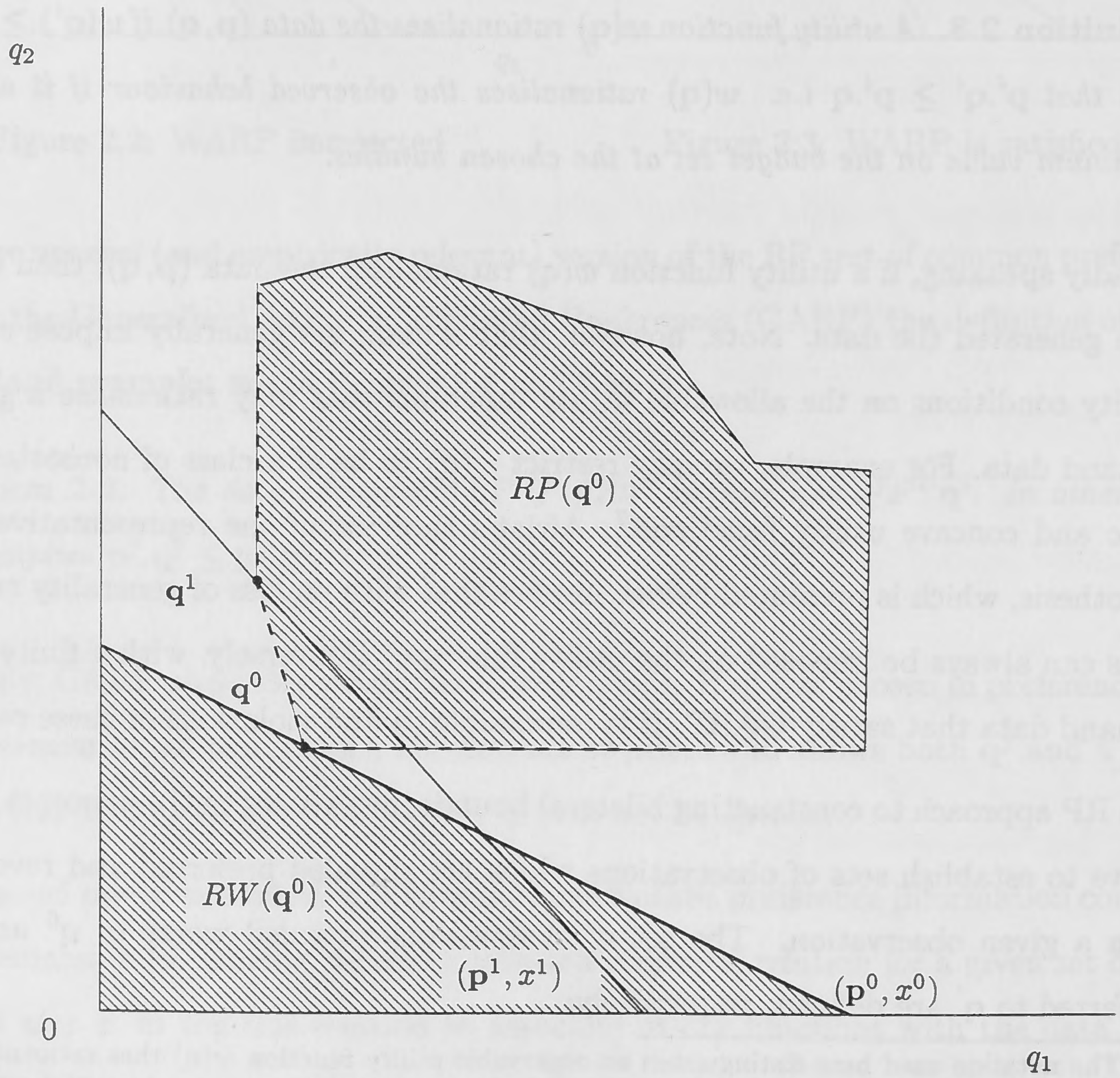

Figure 2.6: Bounding an indifference curve using GARP - example 1

The set of bundles preferred to $\mathbf{q}^0$ (using the true unknown utility function) must always contain $RP(\mathbf{q}^0)$ and must be contained in the complement to $RW(\mathbf{q}^0)$. This set is

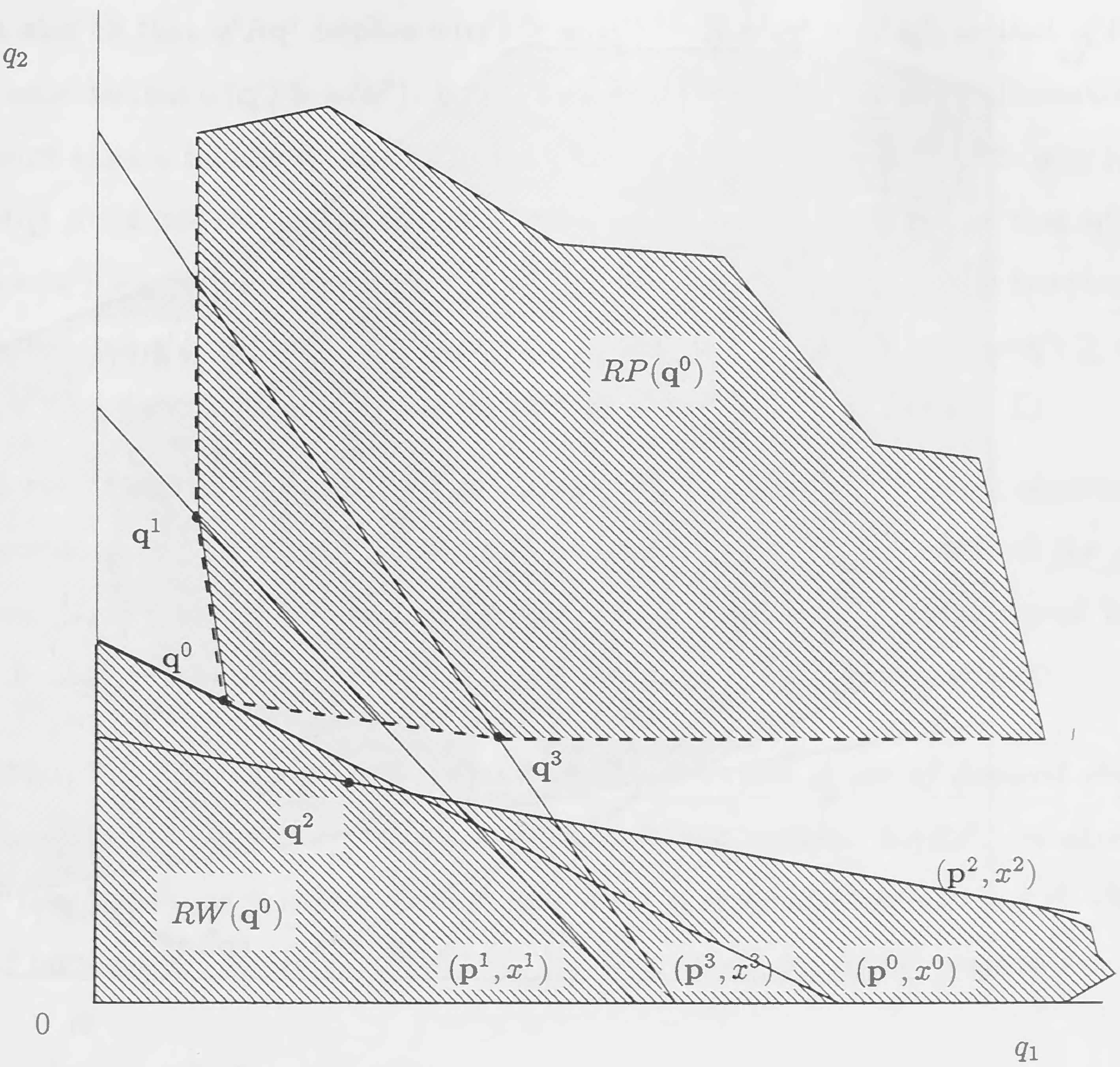

Figure 2.7: Bounding an indifference curve using GARP - example 2

denoted $NRW(\mathbf{q}^0)$, for "not revealed worse" than $\mathbf{q}^0$ and is defined as:

Definition 2.5. $NRW(\mathbf{q}^0) = \{\mathbf{q} : \mathbf{p}^i.\mathbf{q} > \mathbf{p}^i.\mathbf{q}^i$ *for all* $\mathbf{q}^i$ *such that* $\mathbf{q}^0 R\mathbf{q}^i\}$.

$RP(\mathbf{q}^0)$ and $NRW(\mathbf{q}^0)$ thus form "inner" and "outer" estimates of the set of bundles preferred to $\mathbf{q}^0$, and in fact they are the tightest inner and outer estimates (Figure 2.8).

2.2.2 RP tests of common preferences - theory

The RP test of the representative consumer hypothesis is attributable to the work of Afriat (1967), Diewert (1973) and Varian (1982).

Proposition 2.1. Common Preferences - Afriat (1967), Diewert (1973) and Varian (1982) *The following conditions are equivalent:* [9]

[9] In Afriat's original theorem, condition 2 was that the data satisfy "cyclical consistency"; Varian (1982) shows that this is equivalent to the data satisfying the much easier to test GARP.

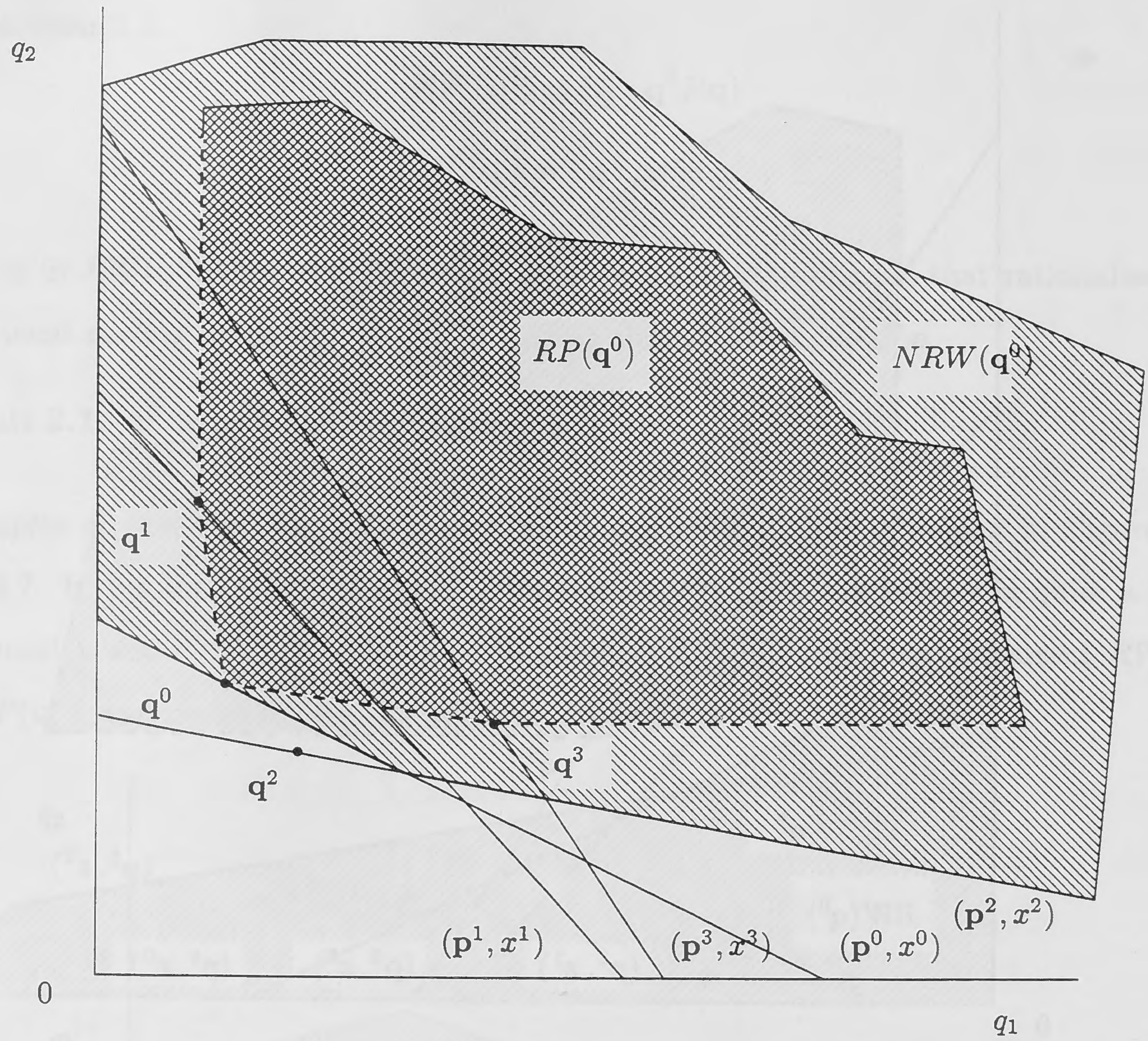

Figure 2.8: $RP(\mathbf{q}^0)$ and $NRW(\mathbf{q}^0)$

(i) There exists a locally non-satiated utility function $w(\mathbf{q})$ that rationalises a finite set of data;

(ii) The data satisfy GARP;

(iii) For all $i, j = 1, ..., N$ there exist numbers W^i, λ^i that satisfy the Afriat inequalities:

(2.1) $$\lambda^i > 0$$

(2.2) $$W^i \leq W^j + \lambda^j \mathbf{p}^j(\mathbf{q}^i - \mathbf{q}^j);$$

(iv) There exists a concave, monotonic, continuous, locally non-satiated utility function $w(\mathbf{q})$ that rationalises the data.

PROOF THAT (i)⇒(ii):[10] Let $w(\mathbf{q})$ rationalise the data. Thus if $\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}^j$, then by definition it must be that $w(\mathbf{q}^i) \geq w(\mathbf{q}^j)$. Therefore $\mathbf{q}^i R^D \mathbf{q}^j$ implies $w(\mathbf{q}^i) \geq w(\mathbf{q}^j)$, and

[10]The following proofs are based on those in Varian (1982) and Diewert and Parkan (1985).

it must also be that $\mathbf{q}^i R \mathbf{q}^j$ implies $w(\mathbf{q}^i) \geq w(\mathbf{q}^j)$.[11] If $\mathbf{p}^i.\mathbf{q}^i > \mathbf{p}^i.\mathbf{q}^j$, so that $\mathbf{q}^i P^D \mathbf{q}^j$, then it must be that $w(\mathbf{q}^i) > w(\mathbf{q}^j)$. If not, then $w(\mathbf{q}^i) = w(\mathbf{q}^j)$, but by local nonsatiation there must exist a bundle $\hat{\mathbf{q}}$ close to $\mathbf{q}^j$ such that $\mathbf{p}^i.\mathbf{q}^i > \mathbf{p}^i.\hat{\mathbf{q}}$ and $w(\hat{\mathbf{q}}) > w(\mathbf{q}^i)$. But then $w(\mathbf{q})$ could not rationalise the data point $(\mathbf{p}^i, \mathbf{q}^i)$. Hence it must be that $\mathbf{q}^i P^D \mathbf{q}^j$ implies $w(\mathbf{q}^i) > w(\mathbf{q}^j)$, and, equivalently, $\mathbf{q}^j P^D \mathbf{q}^i$ implies $w(\mathbf{q}^j) > w(\mathbf{q}^i)$. It has therefore been shown that $\mathbf{q}^i R^D \mathbf{q}^j$ implies not $\mathbf{q}^j P^D \mathbf{q}^i$, since that would mean both $w(\mathbf{q}^i) \geq w(\mathbf{q}^j)$ and $w(\mathbf{q}^j) > w(\mathbf{q}^i)$. Thus GARP follows from $w(\mathbf{q})$ rationalising the data. □

PROOF THAT (ii)⇒(iii): Algorithm 3 in Varian (1982) uses a set of demand observations $(\mathbf{p}, \mathbf{q})$ and a revealed preference relation R that satisfies GARP to construct the Afriat numbers. The proof is conducted by showing that Algorithm 3 (re-produced below) works i.e. that the numbers it produces in fact satisfy the Afriat inequalities.

Algorithm 2.1. CONSTRUCTING AFRIAT NUMBERS *Input: A set of demand observations* $(\mathbf{p}, \mathbf{q})$ *and the revealed preference relation* R *that satisfies GARP. An algorithm* $\max(I)$ *which finds the maximal element from a set of demand observations* I.[12] *Output: A set of numbers* W^i *and* $\lambda^i > 0$, $i = 1, \ldots, N$ *that satisfy the Afriat inequalities.*

1. $I = \{1, \ldots, N\}$, $B = \varnothing$.

2. *Let* $m = \max(I)$.

3. *Set* $E = \{i \text{ in } I : \mathbf{q}^i R \mathbf{q}^m\}$. *If* $B = \varnothing$, *set* $W^i = \lambda^i = 1$ *and go to 6. Otherwise go to 4.*

4. *Set* $W^m = \min_{i \in E} \min_{j \in B} \min\{W^j + \lambda^j \mathbf{p}^j.(\mathbf{q}^i - \mathbf{q}^j), W^j\}$.

5. *Set* $\lambda^m = \max_{i \in E} \max_{j \in B} \max\{(W^j - W^m)/\mathbf{p}^i.(\mathbf{q}^j - \mathbf{q}^i), 1\}$.

6. *Set* $W^i = W^m$, $\lambda^i = \lambda^m$ *for all* $i \in E$.

7. *Set* $I = I \setminus E$, $B = B \cup E$. *If* $I = \varnothing$, *stop. Otherwise go to 2.*

[11] That is, if $\mathbf{q}^i R^D \mathbf{q}^k$, then $w(\mathbf{q}^i) \geq w(\mathbf{q}^k)$ and if $\mathbf{q}^k R^D \mathbf{q}^j$ then $w(\mathbf{q}^k) \geq w(\mathbf{q}^j)$. Hence it must be that $\mathbf{q}^i R \mathbf{q}^j$ implies $w(\mathbf{q}^i) \geq w(\mathbf{q}^j)$.

[12] An element $\mathbf{q}^m$ of a set I is *maximal* with respect to the binary relation R if $\mathbf{q}^i R \mathbf{q}^m$ implies $\mathbf{q}^m R \mathbf{q}^i$ i.e. the only elements ranked ahead of $\mathbf{q}^m$ are those that are indifferent to it. Varian (1982, p.968) provides a two line algorithm, Algorithm 2, for finding the maximal element. Since $(\mathbf{p}, \mathbf{q})$ is a finite set of data and R is a reflexive and transitive binary relations, then there is always at least one maximal element.

At each pass through the algorithm a set of indexes of "equivalent" elements, E, is removed from I and added to B, the set of indexes of "better" elements. The proof that (ii)⇒(iii) involves showing that after step 6 in Algorithm 2.1 is executed, the Ws and λs calculated up to that point all satisfy the Afriat inequalities. Varian (1982, p.969) shows that this is indeed the case.□

PROOF THAT (iii)⇒(iv): From the Afriat inequalities, the function $w(\mathbf{q})$ can be defined:

$$w(\mathbf{q}) = \min_i \{W^i + \lambda^i \mathbf{p}^i.(\mathbf{q} - \mathbf{q}^i)\}$$

This is a piecewise linear function and therefore has the stated properties in (iv). However, it needs to be shown that $w(\mathbf{q})$ rationalises the data. First, note that $w(\mathbf{q}^i) = W^i$ for all $i = 1, \ldots, N$. If this were not the case, and the minimum were in fact obtained at say $\mathbf{q}^m$, then:

$$w(\mathbf{q}^i) = W^m + \lambda^m \mathbf{p}^m.(\mathbf{q}^i - \mathbf{q}^m) \leq W^i + \lambda^i \mathbf{p}^i.(\mathbf{q}^i - \mathbf{q}^i) = W^i.$$

But if this inequality were ever strict, then there would be a violation of one of the Afriat inequalities, and hence it must be that $w(\mathbf{q}^i) = W^i$ for all $i = 1, \ldots, N$.

Suppose there is some $\mathbf{q}$ such that $\mathbf{p}^j\mathbf{q}^j \geq \mathbf{p}^j\mathbf{q}$. For $w(\mathbf{q})$ to rationalise the data, it must be that $w(\mathbf{q}^j) \geq w(\mathbf{q})$. This follows directly from the following set of inequalities:

$$\begin{aligned} w(\mathbf{q}) &= \min_i \{W^i + \lambda^i \mathbf{p}^i.(\mathbf{q} - \mathbf{q}^i)\} \\ &\leq W^j + \lambda^j \mathbf{p}^j.(\mathbf{q} - \mathbf{q}^j) \\ &\leq W^j = w(\mathbf{q}^j), \end{aligned}$$

where the last line is obtained using the fact that $\lambda^j \mathbf{p}^j.(\mathbf{q} - \mathbf{q}^j) \leq 0$ (since $\mathbf{p}^j\mathbf{q}^j \geq \mathbf{p}^j\mathbf{q}$) and $\lambda^j > 0$.□

PROOF THAT (iv)⇒(i): This is obvious.

It has therefore been shown that a test of GARP is equivalent to a test for the existence of common preferences. However, it is worthwhile to also show that (iii)⇐(iv), since this enables an interpretation of the Afriat numbers.

$w(\mathbf{q})$ rationalises the data if it is the case that:

$$\mathbf{q}^i \text{ solves } \max_{\mathbf{q}} [w(\mathbf{q}) : \mathbf{p}^i.\mathbf{q} \leq \mathbf{p}^i.\mathbf{q}^i, \mathbf{q} \geq 0] \qquad i = 1 \ldots N.$$

If it is assumed that $\mathbf{q}^i > 0$ then the non-negativity constraints in the above maximisation problem are not binding.[13] With the further assumption that the utility function is weakly increasing (non-satiation), $w(\mathbf{q})$ rationalises the data if it is the case that:

$$\mathbf{q}^i \text{ solves } \max_{\mathbf{q}}[w(\mathbf{q}) : \mathbf{p}^i.\mathbf{q} = x, \mathbf{q} \geq 0] \qquad i = 1 \ldots N,$$

where x is total expenditure. The Lagrangian for this problem is:

$$L(\mathbf{q}, \lambda) = w(\mathbf{q}) + \lambda(x - \mathbf{p}^i.\mathbf{q}),$$

and the Kuhn-Tucker necessary conditions reduce to: there exist λ^i such that:

$$\begin{aligned} Dw(\mathbf{q}^i) &= \lambda^i \mathbf{p}^i \\ \lambda^i &> 0 \qquad i = 1, ..., N. \end{aligned} \tag{2.3}$$

where $Dw(\mathbf{q}^i) = [\partial w(\mathbf{q}^i)/\partial q_1^i, \ldots, \partial w(\mathbf{q}^i)/\partial q_N^i]^T$. A differentiable concave function has the property that any tangent hyperplane (i.e. any first-order Taylor series expansion) lies above the function. Therefore, for each j, the following inequality is valid for all $\mathbf{q} > 0$:

$$w(\mathbf{q}) \leq w(\mathbf{q}^j) + Dw(\mathbf{q}^j).(\mathbf{q} - \mathbf{q}^j) \qquad j = 1, ..., N.$$

Defining $W^j \equiv w(\mathbf{q}^j)$ for $j = 1, ..., N$ and replacing $\mathbf{q}$ in the above equation with $\mathbf{q}^i, i = 1, ..., N$ gives the following N^2 inequalities:

$$W^i \leq W^j + \lambda^j \mathbf{p}^j.(\mathbf{q}^i - \mathbf{q}^j), \qquad i, j = 1, ..., N. \tag{2.4}$$

Hence, it has been shown that in order for a body of positive data $(\mathbf{p}^i, \mathbf{q}^i) : i = 1, ..., N$ to be rationalised by a concave, monotonic, continuous, non-satiated utility function, it is necessary that there exist $2N$ numbers: $W^1, ..., W^N$ (which can be interpreted as utility levels) and $\lambda^1, ..., \lambda^N$ (which can be interpreted as marginal utilities of income) such that the Afriat inequalities (2.3) and (2.4) are satisfied.□

There are three approaches for testing for common preferences. The first approach described here, testing for GARP, is by far the easiest to implement and is generally used in empirical applications.

[13] Note that in showing that (iii)⇒(iv) it was not necessary to assume that $\mathbf{q}^i > 0$.

2.2.3 Three tests of common preferences

Testing GARP

As GARP is relatively easy to test, the equivalency between (ii) and (iv) in Proposition 2.1 suggests a ready test for common tastes. The first stage of the test of GARP is the construction of the N by N matrix $\mathbf{G}$ which summarises the relation R^D and whose ijth entry is given by:

$$\begin{aligned} G_{ij} &= 1 \qquad \text{if } \mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}^j \text{ i.e. } \mathbf{q}^i R^D \mathbf{q}^j \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

The second stage is the construction of the matrix $\mathbf{H}$ which summarises the relation R:

$$\begin{aligned} H_{ij} &= 1 \qquad \text{if } \mathbf{q}^i R \mathbf{q}^j \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

The matrix $\mathbf{H}$ is constructed by the operation of Warshall's minimum path algorithm (see Varian (1982, Appendix II)), which was originally developed in operations research to calculate the least cost of moving goods from one particular location to another, on $\mathbf{G}$. Warshall's minimum path algorithm is reproduced below (using the Visual Basic programming language).

Algorithm 2.2. WARSHALL'S MINIMUM PATH ALGORITHM *Input:* $G(i,j) = 1$ *if* $\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}^j$, 0 *otherwise. Output:* $H(i,j) = 1$ *if* $\mathbf{q}^i R \mathbf{q}^j$, 0 *otherwise.*

```
For r = 1 To N
    For i = 1 To N
        For j = 1 To N
            If H(i,r) = 1 And H(r,j) = 1 Then H(i,j) = 1
        Next
    Next
Next
```

The matrix $\mathbf{H}$ is used to check for GARP in the following way. If it is the case that $H_{ij} = 1$ and $\mathbf{p}^j.\mathbf{q}^j > \mathbf{p}^j.\mathbf{q}^i$ for some i and j then this is a violation of GARP.

Constructing the Afriat numbers using combinatorial methods

It was shown above that the existence of numbers W^i and λ^i such that the Afriat inequalities (2.1) and (2.2) are satisfied is necessary and sufficient for a given body of data $(\mathbf{p}, \mathbf{q})$ to be consistent with utility maximising behaviour. The second method for checking whether a particular set of data is consistent with common preferences involves calculating the Afriat numbers directly using Algorithm 2.1.

Constructing the Afriat numbers using mathematical programming

The third approach for testing for common preferences is to solve the Afriat inequalities using mathematical programming techniques. Diewert and Parkan (1985) have shown that a necessary and sufficient condition for the existence of numbers W^i and λ^i such that (2.1) and (2.2) are true for a given body of data $(\mathbf{p}, \mathbf{q})$ is that the objective function of the following linear programming problem attain its lower bound of zero:

$$\begin{aligned} &\min_{\boldsymbol{W},\boldsymbol{\lambda},\boldsymbol{S}} \sum_{i\in N}\sum_{j\in N} s^{ij} \quad \text{subject to} \\ &s^{ij} \equiv W^j - W^i - \lambda^i \mathbf{p}^i.(\mathbf{q}^j - \mathbf{q}^i) + S^{ij} \quad i,j \in N \\ &\lambda^i \geq 1 \quad i \in N \\ &s^{ij} \geq 0, S^{ij} \geq 0 \quad i,j \in N, \end{aligned} \tag{2.5}$$

where $\boldsymbol{W} = \{W^1, \ldots, W^N\}$, $\boldsymbol{\lambda} = \{\lambda^1, \ldots, \lambda^N\}$, and s^{ij} and S^{ij} are non-negative slack variables that are used to convert the inequalities (2.2) into equalities. If the above objective function attains it lower bound of zero, then all the optimal $s^{ij} = 0$ and, given $S^{ij} \geq 0$, inequality (2.2) must hold. Obviously $\lambda^i \geq 1$ implies (2.1). Therefore, (2.5) attaining its lower bound of zero is *sufficient* for (2.1) and (2.2) to hold. If the inequalities (2.2) are satisfied with $\lambda^i > 0, i \in N$, then all of the numbers W^i and λ^i can be scaled so that the same inequalities (2.2) are satisfied with $\lambda^i \geq 1, i \in N$. Hence the linear programming problem (2.5) will attain its lower bound of zero. It has therefore been shown that (2.5) attaining its lower bound of zero is both necessary and sufficient for set of demand observations $(\mathbf{p}, \mathbf{q})$ to be rationalised by a utility function and hence for the existence of common preferences.

The constrained linear programming problem (2.5) can be converted into the following unconstrained non-linear programming problem:

$$\min_{W,\alpha,\sigma} \sum_{i\in N}\sum_{j\in N}[s^{ij}]^2 \quad \text{subject to}$$

$$s^{ij} \equiv U^j - U^i - (\alpha_i^2+1)\mathbf{p}^i.(\mathbf{q}^j - \mathbf{q}^i) + \sigma_{ij}^2 \qquad i,j \in N. \tag{2.6}$$

If we define $\lambda^i \equiv \alpha_i^2 + 1 \geq 1$ and $S^{ij} \equiv \sigma_{ij}^2 \geq 0$ then it is apparent that (2.6) corresponds to (2.5).

The above tests of common non-homothetic preferences may also be conducted using normalised prices $\mathbf{v}^i = \mathbf{p}^i/\mathbf{p}^i.\mathbf{q}^i$. Replacing $\mathbf{p}^j$ in (2.2) with $\mathbf{v}^j$ gives:

$$W^i \leq W^j + \lambda^j(\mathbf{v}^j.\mathbf{q}^i - 1) \qquad i,j \in N. \tag{2.7}$$

Using normalised prices, the constrained and unconstrained programming problems for the tests of common non-homothetic preferences are respectively:

$$\begin{aligned} &\min_{W,\lambda,S} \sum_{i\in N}\sum_{j\in N} s^{ij} \quad \text{subject to} \\ &s^{ij} \equiv W^j - W^i - \lambda^i(\mathbf{v}^i.\mathbf{q}^j - 1) + S^{ij} \qquad i,j \in N \\ &\lambda^i \geq 0 \qquad i \in N \\ &s^{ij} \geq 0, S^{ij} \geq 0 \qquad i,j \in N. \end{aligned} \tag{2.8}$$

$$\begin{aligned} &\min_{W,\alpha,\sigma} \sum_{i\in N}\sum_{j\in N}[s^{ij}]^2 \quad \text{subject to} \\ &s^{ij} \equiv W^j - W^i - (\alpha_i^2+1)(\mathbf{v}^i.\mathbf{q}^j - 1) + \sigma_{ij}^2 \qquad i,j \in N. \end{aligned} \tag{2.9}$$

2.2.4 Example

The three RP tests for common general preferences were applied to example data for 6 countries (note that in practice, only the test of GARP would be used as it is much easier to implement). The price vectors are: $\mathbf{p}^1 = \{3,1\}$, $\mathbf{p}^2 = \{0.5,1\}$, $\mathbf{p}^3 = \{1,1\}$, $\mathbf{p}^4 = \{5,1\}$, $\mathbf{p}^5 = \{0.25,1\}$ and $\mathbf{p}^6 = \{1/3,1\}$, while the commodity bundles are: $\mathbf{q}^1 = \{1.25, 2.25\}$, $\mathbf{q}^2 = \{7,2.5\}$, $\mathbf{q}^3 = \{2.5,2.5\}$, $\mathbf{q}^4 = \{1.5,7.5\}$, $\mathbf{q}^5 = \{6,0.5\}$ and $\mathbf{q}^6 = \{4.5,2.5\}$. The data are shown in Figure 2.9.

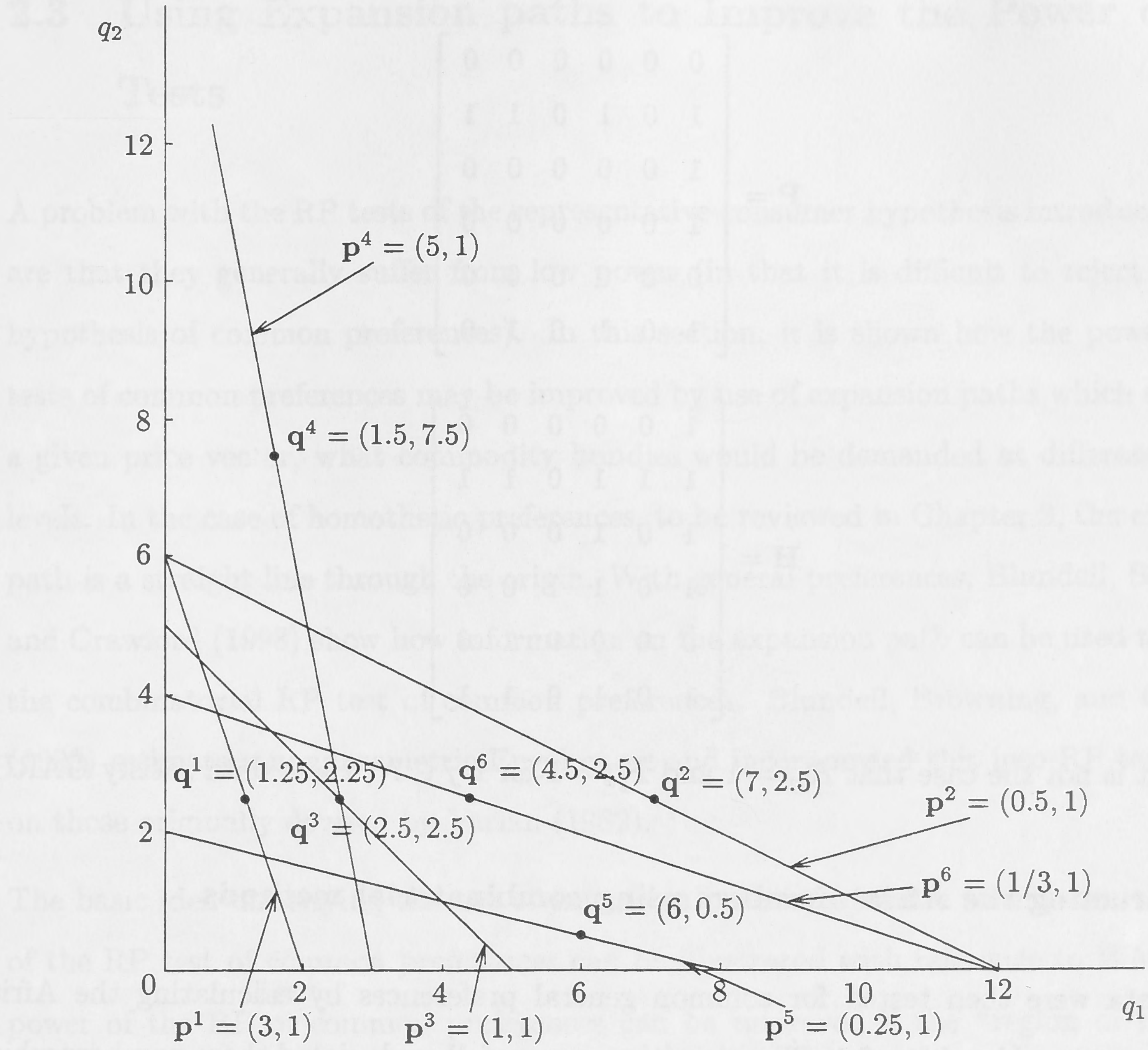

Figure 2.9: Six country demand data

Test of GARP

For these data, the matrices relevant to the test of GARP are (where $P_{ij} = 1$ if $\mathbf{q}^i P^D \mathbf{q}^j$ and $P_{ij} = 0$, otherwise):

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Since it is not the case that $H_{ij} = 1$ and $P_{ji} = 1$ for $i, j \in N$, these data satisfy GARP.

Constructing the Afriat numbers using combinatorial methods

The data were then tested for common general preferences by calculating the Afriat numbers using Algorithm 2.1. The algorithm successfully calculated these numbers which are: $W^1 = -5.5$, $W^2 = 1$, $W^3 = -1.25$, $W^4 = 1$, $W^5 = -5.5$, $W^6 = -0.25$; $\lambda^1 = 1.083$, $\lambda^2 = 1$, $\lambda^3 = 1$, $\lambda^4 = 1$, $\lambda^5 = 3.778$ and $\lambda^5 = 1.5$. Note, however that these numbers are not unique (there will, in general, be an infinite combination of numbers which satisfy the Afriat inequalities), and hence they cannot be used for computing utility bounds.

Constructing the Afriat numbers using mathematical programming

The third RP test of the existence of a representative consumer involves solving the mathematical program (2.9) for these data. For the particular starting values used, and normalising $W^3 = 0$, the Afriat numbers calculated are: $W^1 = -0.663$, $W^2 = 0.750$, $W^3 = 0$, $W^4 = 0.341$, $W^5 = -0.268$, $W^6 = 0.431$; $\lambda^1 = 1.005$, $\lambda^2 = 1.362$, $\lambda^3 = 1.078$, $\lambda^4 = 1.717$, $\lambda^5 = 1.445$ and $\lambda^6 = 1.529$. As above, however, these Afriat numbers are not unique and hence do not provide any information on the bounds to the bilateral true welfare indexes.

2.3 Using Expansion paths to Improve the Power of RP Tests

A problem with the RP tests of the representative consumer hypothesis introduced above are that they generally suffer from low power (in that it is difficult to reject the null hypothesis of common preferences). In this section, it is shown how the power of RP tests of common preferences may be improved by use of expansion paths which show, for a given price vector, what commodity bundles would be demanded at different income levels. In the case of homothetic preferences, to be reviewed in Chapter 3, the expansion path is a straight line through the origin. With general preferences, Blundell, Browning, and Crawford (1998) show how information on the expansion path can be used to modify the combinatorial RP test of common preferences. Blundell, Browning, and Crawford (1998) estimate a nonparametric Engel curve and incorporated this into RP tests based on those originally devised by Varian (1982).[14]

The basic idea underlying the use of expansion path information to improve the power of the RP test of common preferences can be illustrated with reference to WARP. The power of the RP test of common preferences can be improved if the "region of rejection" (the segment of a budget hyperplane where a consumption bundle must lie for revealed preference to be rejected) can be increased. In Figure 2.10, given $\mathbf{q}^0$ has been observed, $\mathbf{q}^1$ can be anywhere along the budget hyperplane $(\mathbf{p}^1, x^1)$, and there will be no rejection of revealed preference, i.e. there is no region of rejection. Thus in this particular example, the standard WARP test of common preferences has zero power (and, in general, WARP has low power).

With knowledge of the expansion path $EP(\mathbf{p}^0)$, $\mathbf{q}^{0^*}$ is revealed preferred to bundles in the shaded region and the bold segment of $(\mathbf{p}^1, x^1)$ therefore becomes the region of rejection; if $\mathbf{q}^1$ is on the bold section of $(\mathbf{p}^1, x^1)$ then we have $\mathbf{q}^1 R^D \mathbf{q}^{0^*}$ and $\mathbf{q}^{0*} R^D \mathbf{q}^1$ and hence a violation of WARP. Therefore, for WARP to hold, $\mathbf{q}^1$ must be on the segment of $(\mathbf{p}^1, x^1)$ which is not bold.

If the budget hyperplanes do not intersect, then the standard WARP test of common

[14]The mathematical programming approach does not easily lend itself to incorporating expansion path information to improve the power of the test of common preferences, so it is not discussed further in this context. However, Chavas and Cox (1997) show how bounds on expansion paths can be used in the construction of true indexes using the envelope functions suggested by Afriat (1987).

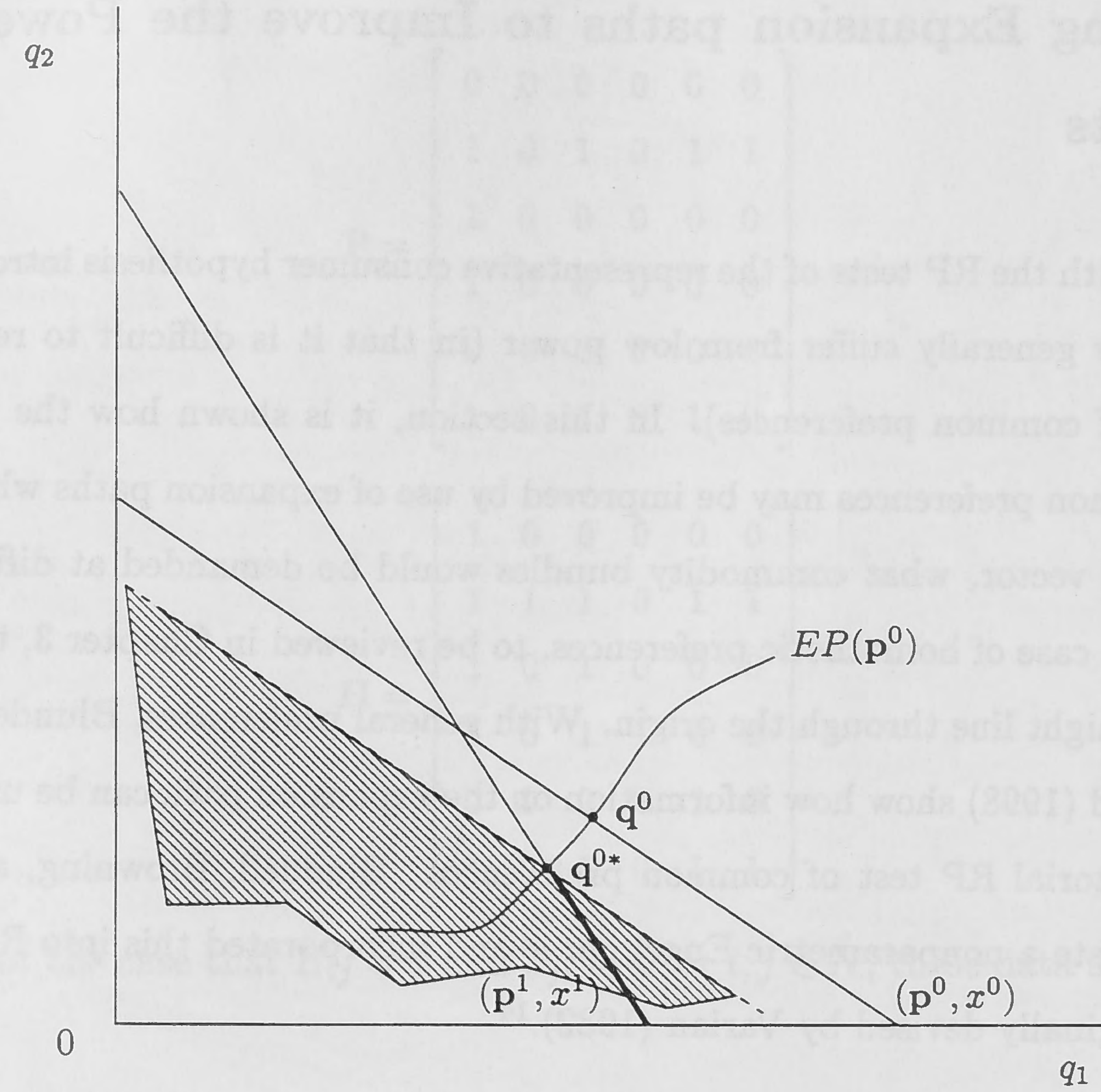

Figure 2.10: Using the expansion path to improve the power of WARP

preferences (where there is no knowledge of the expansion paths) will always have zero power (i.e. regardless of where $\mathbf{q}^0$ may be on the budget line, WARP will always be satisfied) . However, even when the budget hyperplanes do not intersect, if the expansion path is known then the power of the WARP test can be improved in a manner similar to that illustrated in Figure 2.10.

2.3.1 Sequential maximum power path

The more general (and empirically relevant) version of the RP test of common preferences is the test of GARP. In this sub section, the approach of Blundell, Browning, and Crawford (1998) for using expansion paths to improve the power of the GARP test of common preferences is reviewed. Recall the $(N \times N)$ matrix $\mathbf{G}$ introduced above which is constructed so that $G_{st} = 1$ if $\mathbf{q}^s(x^s)R^D\mathbf{q}^t(x^t)$.[15] The sub-sequence of total expenditures $\{x^s, x^t, \ldots, x^v, x^w\}$ is said to be *preference ordered* if $\{G_{st}, G_{tu}, \ldots, G_{vw}\} = \{1, 1, \ldots, 1\}$ i.e. if the demand associated with any total outlay is revealed at least as good as the next

[15]For notational convenience $\mathbf{p}$ is suppressed in the expression for the Marshallian demand.

one. From this, $\mathbf{q}^s(x^s)R\mathbf{q}^w(x^w)$ if there is a preference-ordered sub-sequence starting in s and ending in w. If, for a given preference-ordered sub-sequence $\{x^s, x^t, ..., x^v, x^w\}$, it is found that $\mathbf{q}^w(x^w)P^D\mathbf{q}^s(x^s)$, then the sub-sequence fails GARP.

For a given sequence of demands, consider any preference-ordered sub-sequence $\{x^s, x^t, \dots, x^v, x^w\}$. Given total expenditure in the last country in the sub-sequence, x^w, total outlay in the second-to-last country v is chosen so that the country w bundle is just affordable at country v prices; denote this $\tilde{x}^v = \mathbf{p}^v.\mathbf{q}^w(x^w)$. Thus $\mathbf{q}^v(\tilde{x}^v)$ is directly revealed preferred to $\mathbf{q}^w(x^w)$. Total outlay in the previous country is then chosen so that $\mathbf{q}^v(\tilde{x}^v)$ is just affordable and so on. The *sequential maximum power* (SMP) path for the preference ordered sub-sequence $\{x^s, x^t, ..., x^v, x^w\}$ is given by:

$$\tilde{x}^s, \tilde{x}^t, ..., \tilde{x}^v, \tilde{x}^w = \mathbf{p}^s.\mathbf{q}^t(\tilde{x}^t), \mathbf{p}^t.\mathbf{q}^u(\tilde{x}^u), ..., \mathbf{p}^v.\mathbf{q}^w(\tilde{x}^w), x^w.$$

By construction, any SMP path is preference ordered. In the three country, two good example of Figure 2.11, the order of the preference-ordered sub-sequence is $\mathbf{q}^2R^D\mathbf{q}^1R^D\mathbf{q}^0$. GARP is rejected (since $\mathbf{q}^2R\mathbf{q}^0$ and $\mathbf{q}^0R^D\mathbf{q}^2$); however, if $\mathbf{q}^2$ is on the segment of $(\mathbf{p}^2, x^2)$ which is between $\mathbf{q}^1$ and the bold segment, then GARP will be accepted. With normal demands, it is apparent that if $(\mathbf{p}^2, x^2)$ is "pushed out", the region of rejection (the bold segment of the budget hyperplane) will become shorter relative to the region where GARP will be accepted. This will minimise the probability of finding some rejection of GARP, and hence the power of the RP test of common preferences will decrease.

In Figure 2.12, hypothetical expansion paths are added to the example of Figure 2.11. To construct the SMP path, start at the budget line $(\mathbf{p}^0, x^0)$ at which $\mathbf{q}^0(x^0)$ is demanded. Choose a budget line under $\mathbf{p}^1$ so that $\mathbf{q}^0$ can just be bought at $\mathbf{p}^1$; this is labeled $(\mathbf{p}^1, \tilde{x}^1)$. The chosen $\mathbf{q}^1(\tilde{x}^1)$ is such that $\mathbf{q}^1(\tilde{x}^1)R^D\mathbf{q}^0(\tilde{x}^0)$ (where the notation $\tilde{x}^0$ denotes that x^0 is a member of a preference ordered sub-sequence that rejects GARP) and it is found by moving along country 1's expansion path $EP(\mathbf{p}^1)$. Since the observed budget line was $(\mathbf{p}^1, x^1)$, this is an example in which the budget line has shifted out, thus reducing the probability of finding a rejection. Now choose a budget line under $\mathbf{p}^2$ so that $\mathbf{q}^1$ can just be bought at $\mathbf{p}^2$; this is labeled $(\mathbf{p}^2, \tilde{x}^2)$. Now we have $\mathbf{q}^2(\tilde{x}^2)R^D\mathbf{q}^1(\tilde{x}^1)R^D\mathbf{q}^0(\tilde{x}^0)$ and the region of rejection of $\mathbf{q}^2$ with respect to $\mathbf{q}^0$ has been maximised.

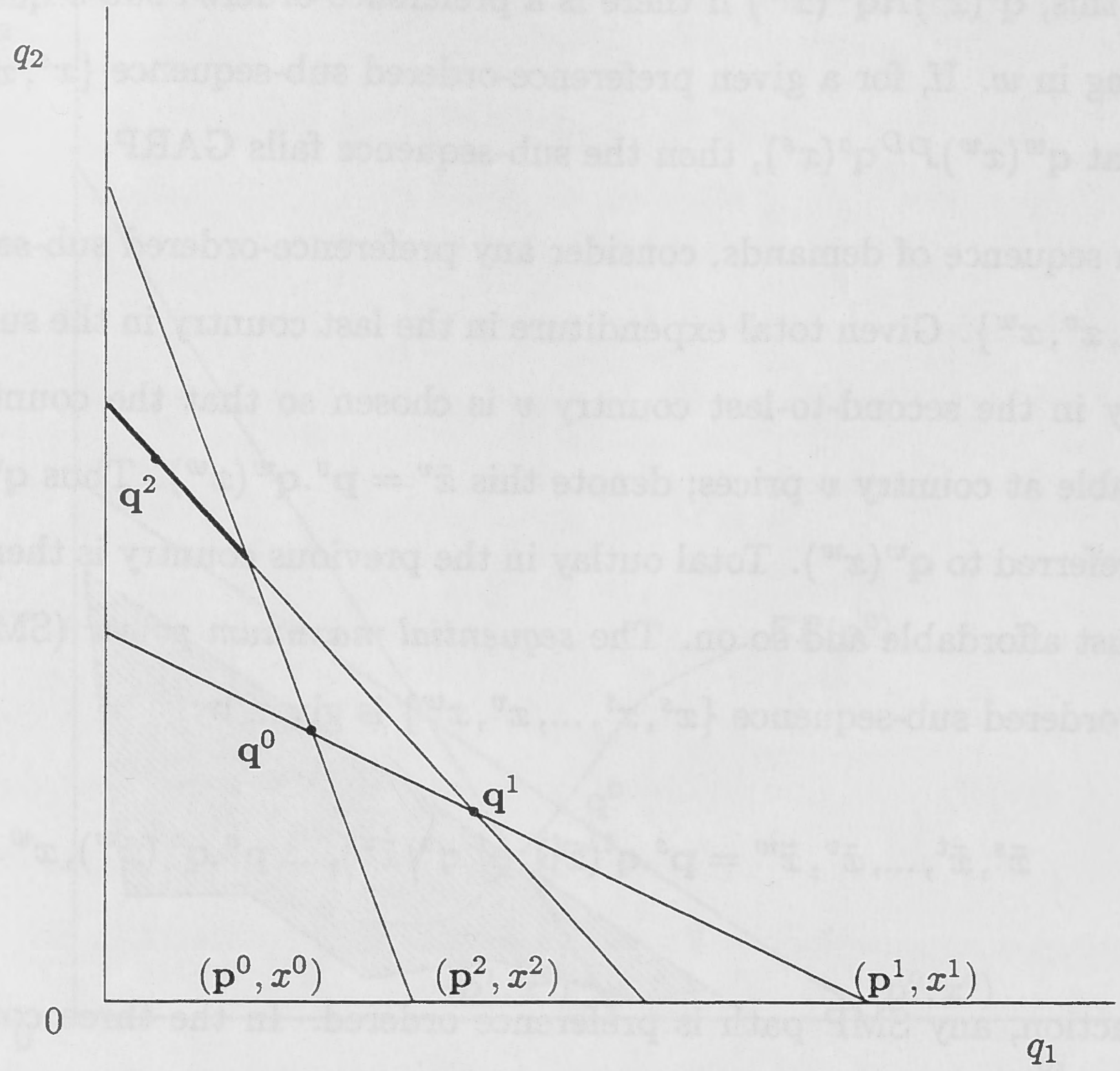

Figure 2.11: A preference-ordered sub-sequence which fails GARP

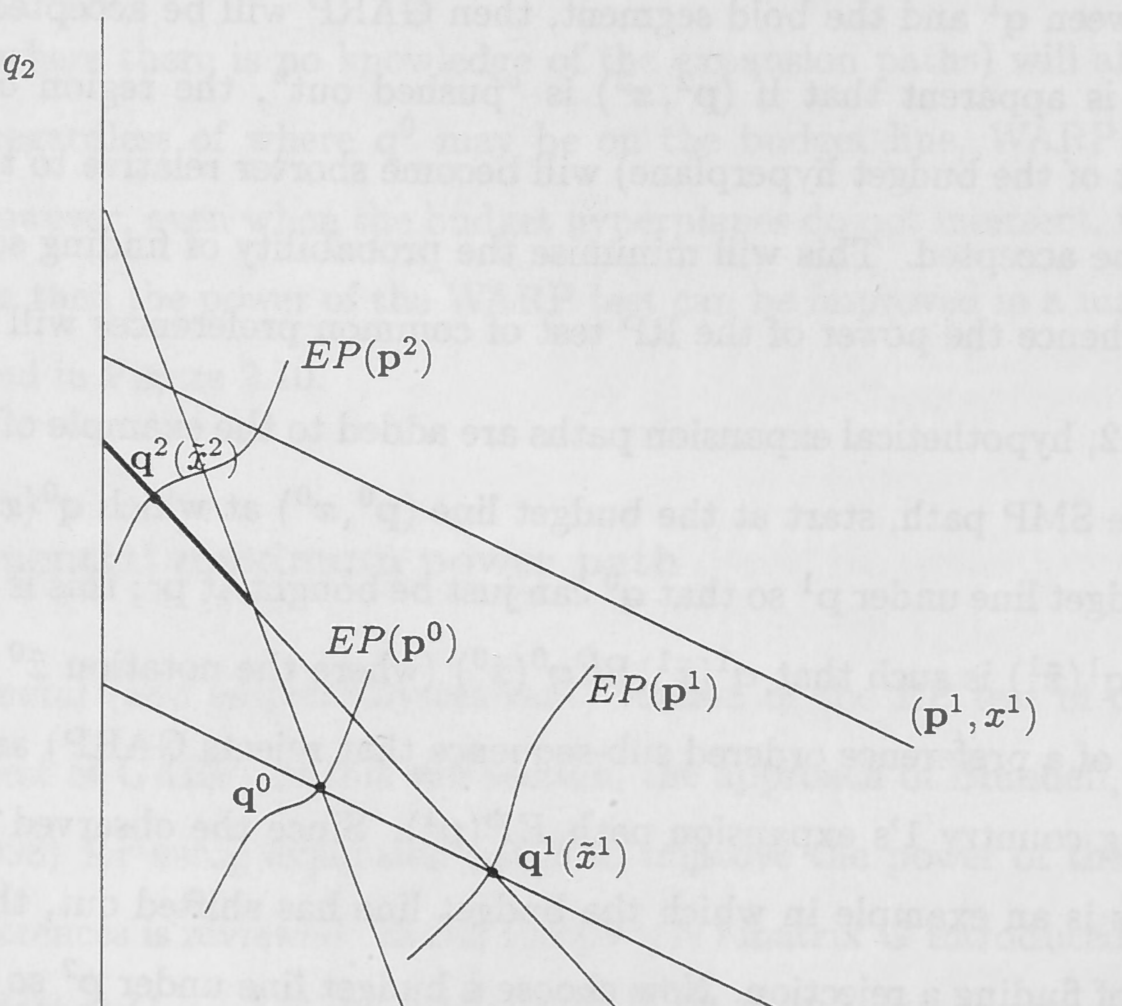

Figure 2.12: Increasing the power of GARP using the SMP path

2.3.2 Bounding indifference curves using the BBC algorithms

Blundell, Browning, and Crawford (1998) construct algorithms which operationalise their more powerful RP test of common preferences. The BBC algorithms provide upper- and lower-bounds on an indifference curve through a given point in commodity space; if these bounds can be constructed for all countries in the data set, then the data satisfy common preferences.

Algorithm 2.3. $RP(\mathbf{q}^0)$ BOUND ALGORITHM - BLUNDELL, BROWNING, AND CRAWFORD (1998) *Input: A set of demand observations* $(\mathbf{p}, \mathbf{q})$ *and estimated Engel curve. Output: The set RP of boundary points of which* $\mathbf{q}^0$ *is a member and which has* $N+1$ *elements where* $\mathbf{p}^i.\mathbf{q}^i \le \mathbf{p}^i.\mathbf{q}^j \ \forall\ \mathbf{q}^i, \mathbf{q}^j \in RP$ *and either* $\mathbf{q}^i R^D \mathbf{q}^0$ *or* $\mathbf{q}^j R^D \mathbf{q}^0 \ \forall\ \mathbf{q}^i \in RP$.

1. *Set* $W = \{\mathbf{q}^0\},\ \ \tau = \{0, 1, \ldots, N\},\ \ E = \emptyset$.
2. *Set* $R = \{\mathbf{q}^t = (\min\{x|\mathbf{p}^t.\mathbf{q}^t = \mathbf{p}^t.\mathbf{q}^w\})\ \ \mathbf{q}^w \in W,\ \ t \in \tau\}$.
3. *Set* $E = \{\mathbf{q}^i \in R : \mathbf{p}^i.\mathbf{q}^i > \mathbf{p}^i.\mathbf{q}^j \text{ for } \mathbf{q}^j \in R\}$.
4. *Set* $W = R/E$.
5. *if* $E = \emptyset$ *set* $RP = W$ *and stop. Otherwise go to (2).*

Algorithm 2.4. $RW(\mathbf{q}^0)$ BOUND ALGORITHM - BLUNDELL, BROWNING, AND CRAWFORD (1998) *Input: A set of demand observations* $(\mathbf{p}, \mathbf{q})$ *and estimated Engel curve. Output: The set RW of boundary points of which* $\mathbf{q}^0$ *is a member and where* $\mathbf{p}^i.\mathbf{q}^i \le \mathbf{p}^i.\mathbf{q}^j$ $\forall\ \mathbf{q}^i, \mathbf{q}^j \in RW$ *and either* $\mathbf{q}^0 R^D \mathbf{q}^i$ *or* $\mathbf{q}^0 R^D \mathbf{q}^j \ \forall\ \mathbf{q}^i \in RP$.

1. *Set* $B = \{\mathbf{q}^0\},\ \ \tau = \{0, 1, \ldots, N\},\ \ E = \emptyset$.
2. *Set* $R = \{\mathbf{q}^t = (\max\{x|\mathbf{p}^b.\mathbf{q}^b = \mathbf{p}^b.\mathbf{q}^t\})\ \ \mathbf{q}^b \in B,\ \ t \in \tau\}$.
3. *Set* $E = \{\mathbf{q}^i \in R : \mathbf{p}^j.\mathbf{q}^j > \mathbf{p}^j.\mathbf{q}^i \text{ for } \mathbf{q}^j \in R\}$.
4. *Set* $B = R/E$.
5. *if* $E = \emptyset$ *set* $RW = B$ *and stop. Otherwise go to (2).*

Blundell, Browning, and Crawford (1998, p.20) prove that if the data local to the reference bundle $\mathbf{q}^0$ reject GARP, then $RP(\mathbf{q}^0)$ and $RW(\mathbf{q}^0)$ Bound Algorithms will not converge.

It is now illustrated how the BBC algorithms compute the upper bound to the indifference curve passing through $\mathbf{q}^0$ (see Figure 2.13). The process involves iteratively constructing three sets of bundles. Set W contains all bundles on the upper-bound to the indifference curve passing $\mathbf{q}^0$ (initially W only contains $\mathbf{q}^0$). Set R contains the bundles found to be revealed preferred to the bundles in W in any particular iteration. Set E contains those bundles in R which are revealed preferred to at least one other bundle in R. In every iteration, W is updated as $W = R/E$ (i.e. W contains those bundles which are on the boundary of the set of bundles revealed preferred to $\mathbf{q}^0$, denoted $RP(\mathbf{q}^0)$).

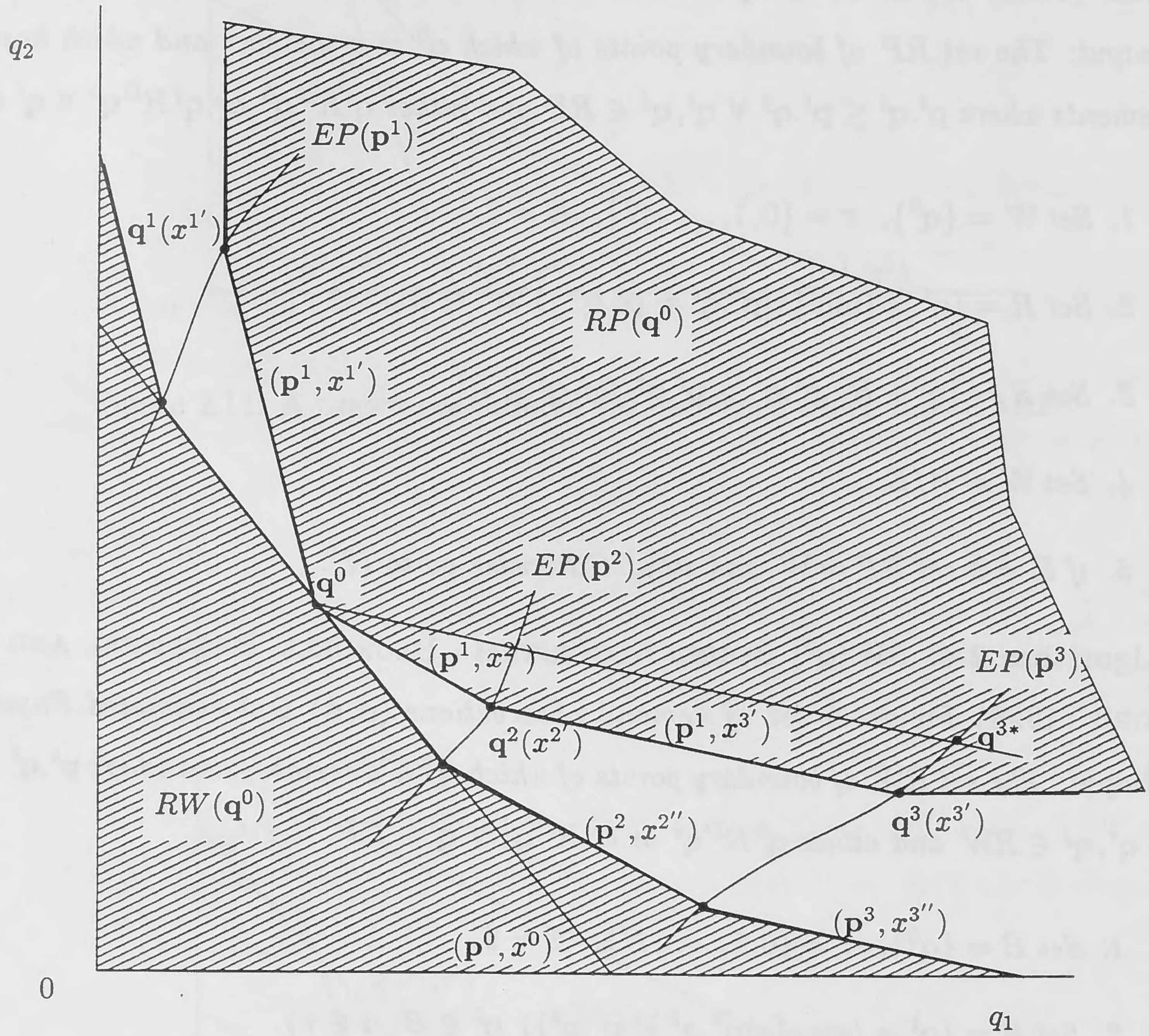

Figure 2.13: Bounding an indifference curve using the BBC algorithms

In this example $\mathbf{q}^1(x^{1'})$ and $\mathbf{q}^2(x^{2'})$ are both directly preferred to $\mathbf{q}^0$ and hence are added to W in the first iteration. The point $\mathbf{q}^{3*}$ will also be added to R in the first iteration (since $\mathbf{p}^3.\mathbf{q}^{3*} = \mathbf{p}^3.\mathbf{q}^0$), but since $\mathbf{q}^{3*}P^D\mathbf{q}^2(x^{2'})$, $\mathbf{q}^{3*}$ lies above the indifference curve and will be added to E in the first iteration (and therefore excluded from W in the first iteration). At the end of the first iteration, we therefore have three bundles in

$W : \mathbf{q}^0, \mathbf{q}^1(x^{1'})$ and $\mathbf{q}^2(x^{2'})$.

In the second iteration R will contain the bundles which are revealed preferred to each of the three current members of W. The bundles revealed preferred to $\mathbf{q}^0$ will be $\mathbf{q}^0$, $\mathbf{q}^1(x^{1'})$, $\mathbf{q}^2(x^{2'})$ and $\mathbf{q}^{3*}$ again. There will be another four bundles (one on each of the four expansion paths including $EP(\mathbf{p}^0)$, which is not shown in the figure) which are revealed preferred to each of the other members of W $\{\mathbf{q}^1(x^{1'}), \mathbf{q}^2(x^{2'})\}$ but, except for $\mathbf{q}^3(x^{3'})$, these will all go into E since step 2 of Algorithm 2.3 selects the cheapest bundle on each expansion path. None of the four bundles in R will be strictly preferred to the others, and hence W will contain four bundles: $\mathbf{q}^0, \mathbf{q}^1(x^{1'}), \mathbf{q}^2(x^{2'})$ and $\mathbf{q}^3(x^{3'})$. The newest addition to W, $\mathbf{q}^3(x^{3'})$, is chosen such that $\mathbf{p}^3.\mathbf{q}^3(x^{3'}) = \mathbf{p}^3.\mathbf{q}^2(x^{2'})$. Therefore, $\mathbf{q}^3(x^{3'})R^D\mathbf{q}^2(x^{2'})$ and $\mathbf{q}^2(x^{2'})R^D\mathbf{q}^0$ and hence $\mathbf{q}^3(x^{3'})R\mathbf{q}^0$. Since $E = \emptyset$, the algorithm terminates and the upper bound to the indifference curve passing through $\mathbf{q}^0$ has thus been determined (the lower bound is constructed in an analogous manner).

2.3.3 Example

The improved GARP test of the representative consumer hypothesis is now illustrated for the 6 country data. Assume that the following Engel curves have been estimated for the data (note it is assumed for simplicity that preferences are quasi-homothetic, and hence the Engel curves are linear):

$$
\begin{array}{ll}
q_1^1 = -1/4 + 1/4x & q_2^1 = 3/4 + 1/4x \\
q_1^2 = 7/13 + 14/13x & q_2^2 = -7/26 + 6/13x \\
q_1^3 = x/2 & q_2^3 = x/2 \\
q_1^4 = -33/58 + 8/58x & q_2^4 = 165/58 + 18/58x \\
q_1^5 = 6/5 + 12/5x & q_2^5 = -3/10 + 2/5x \\
q_1^6 = 9/10 + 9/10x & q_2^6 = -3/10 + 7/10x.
\end{array}
$$

The expansion paths implied by these Engel curves are shown in Figure 2.14.

Algorithms 2.3 and 2.4 were run for the 6 country data using the above Engel curves, and these algorithms would not converge thus indicating that the data do not satisfy the improved RP test of common preferences. The reason the data do not satisfy the improved test of GARP be illustrated by considering the construction of the RP bound to the indifference curve passing through $\mathbf{q}^3$ (Figures 2.15 and 2.16).

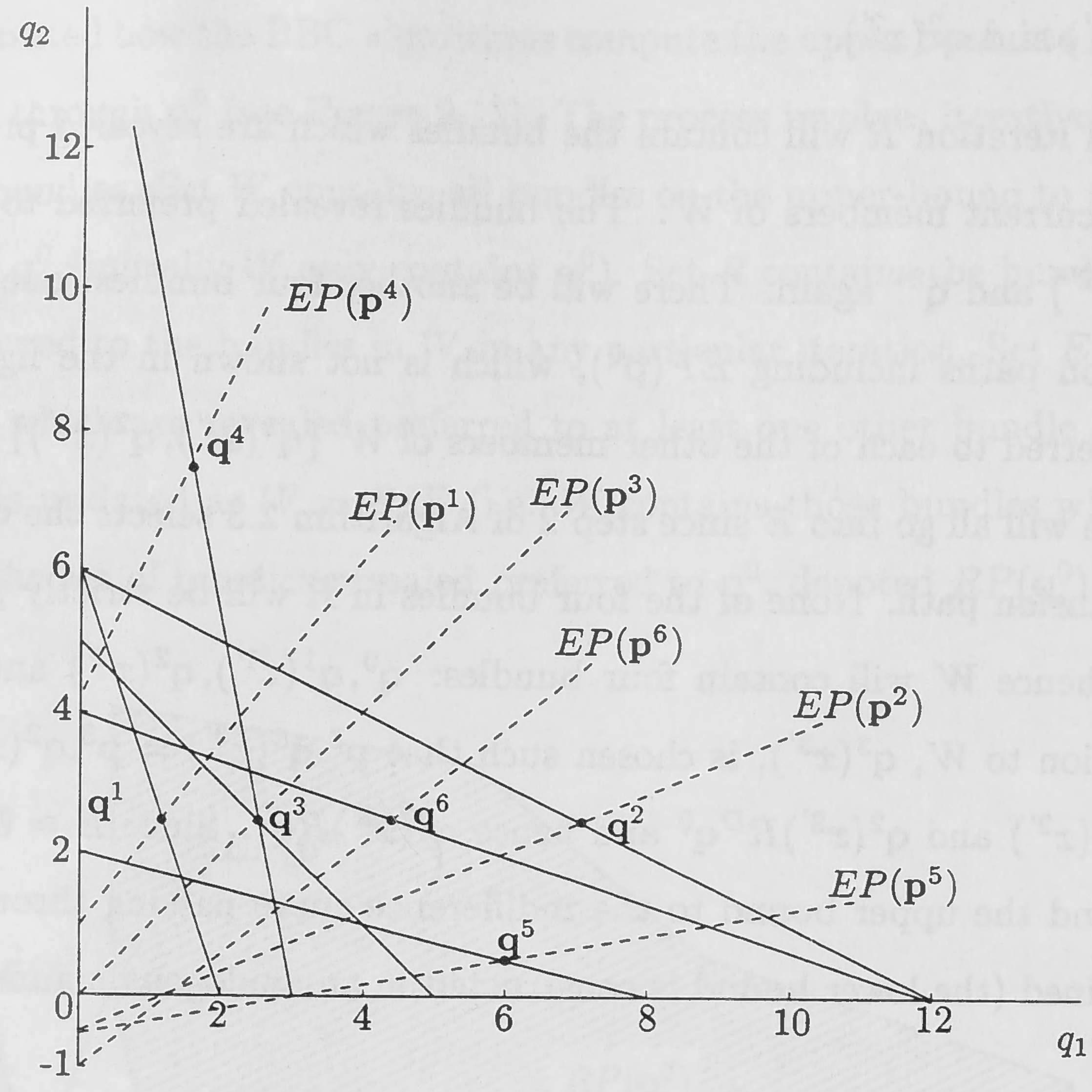

Figure 2.14: Quasi-homothetic expansion paths

In the first iteration of the $RP(\mathbf{q}^3)$ bound algorithm, W contains three bundles $W = \{\mathbf{q}^3, \mathbf{q}^{1*}, \mathbf{q}^{2*}\}$. The bundles on the other three expansion paths ($\mathbf{q}^{4*}$, $\mathbf{q}^{5*}$, $\mathbf{q}^{6*}$) are not included in W in the first iteration since $\mathbf{q}^{4*}P^D\mathbf{q}^{1*}$, $\mathbf{q}^{6*}P^D\mathbf{q}^{2*}$ and $\mathbf{q}^{5*}P^D\mathbf{q}^{2*}$ and hence in the first iteration $E = \{\mathbf{q}^{4*}, \mathbf{q}^{6*}, \mathbf{q}^{5*}\}$. However, in the second iteration, $W = \{\mathbf{q}^3, \mathbf{q}^{1*}, \mathbf{q}^{4**}, \mathbf{q}^{6**}, \mathbf{q}^{5**}\}$ and $E = \{\mathbf{q}^{2*}\}$ (since $\mathbf{q}^{2*}P^D\mathbf{q}^{6**}$). Thus, $\mathbf{q}^{2*}$ has been dropped from W and $RP(\mathbf{q}^3)$ will not converge since in the next iteration a bundle on $EP(\mathbf{p}^2)$ will replace $\mathbf{q}^{6**}$ and $\mathbf{q}^{5**}$ in W, and so on. With country 6 dropped, however, the data satisfy the improved GARP test of common preferences. Note that the complete bounds to the indifference curve passing through $\mathbf{q}^3$ is shown in Figure 2.17.

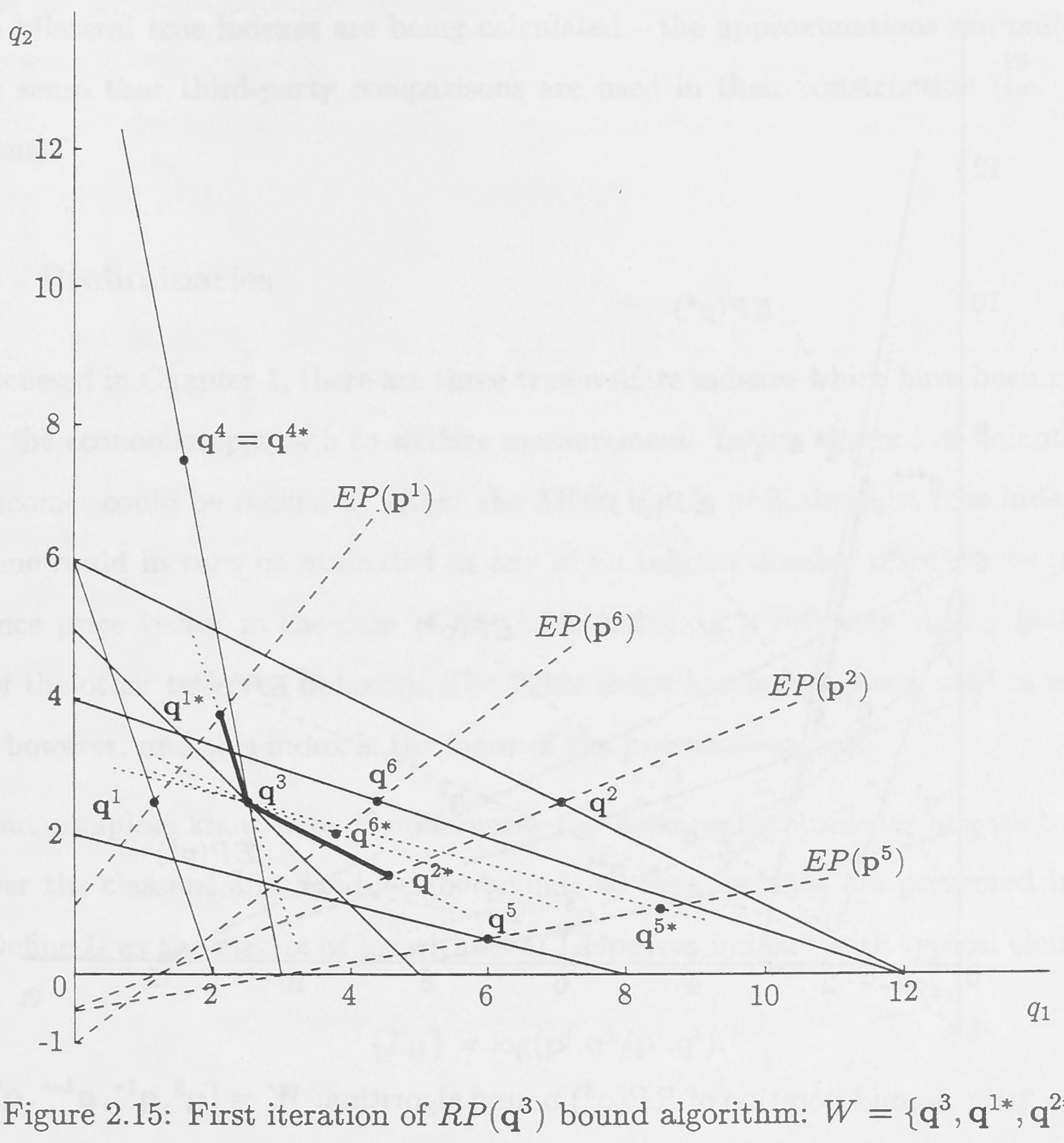

Figure 2.15: First iteration of $RP(\mathbf{q}^3)$ bound algorithm: $W = \{\mathbf{q}^3, \mathbf{q}^{1*}, \mathbf{q}^{2*}\}$

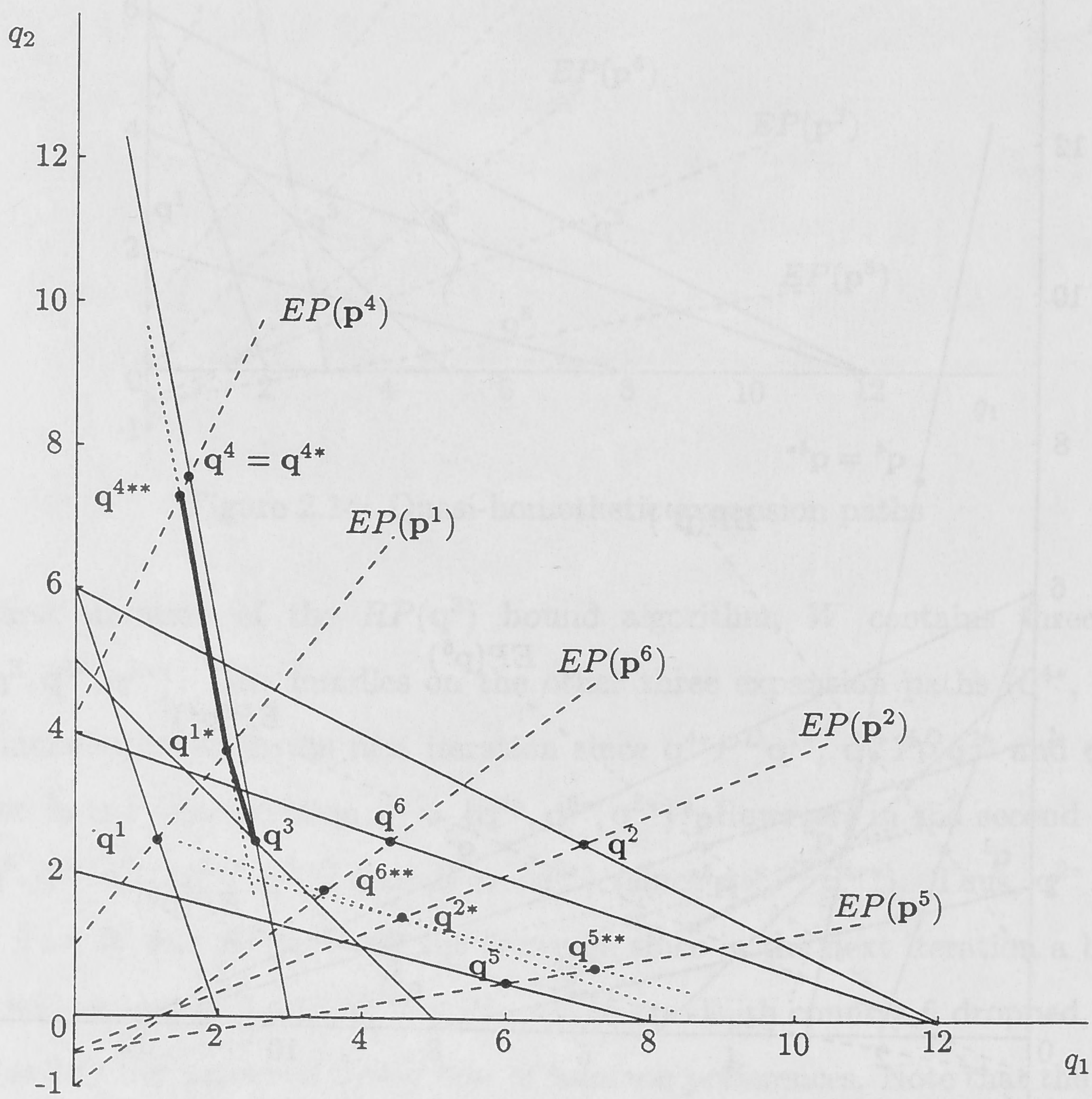

Figure 2.16: Second iteration of $RP(\mathbf{q}^3)$ bound algorithm: $W = \{\mathbf{q}^3, \mathbf{q}^{1*}, \mathbf{q}^{4**}, \mathbf{q}^{6**}, \mathbf{q}^{5**}\}$

2.4 Constructing Bilateral True Indexes

Under the RP approach to welfare measurement, bilateral true indexes of welfare can only be constructed when the data do not reject common preferences. The approach used is first to test for common preferences and then for those countries which share common preferences, the bounds to the bilateral true indexes (which, as argued above, can be regarded as true indexes themselves) are constructed. Thus, multilateral approximations to the bilateral true indexes are being calculated - the approximations are multilateral in the sense that third-party comparisons are used in their construction (i.e. via RP relations).

2.4.1 Preliminaries

As discussed in Chapter 1, there are three true welfare indexes which have been proposed under the economic approach to welfare measurement. In the absence of homotheticity, real incomes could be ranked by either the Allen, Konüs or Malmquist true indexes, and each one could in turn be evaluated at any of an infinite number of reference points (a reference price vector in the case of the Allen index, or a reference utility level in the case of the other two true indexes). The Allen index has mainly been used in empirical work, however, and this index is the focus of the presentation here.

Without complete knowledge of preferences the Allen welfare indexes are unobservable, however the classical and fixed-weight bounds to these indexes are presented in Result 1.2. Define $\mathbf{L}$ as the matrix of logarithms of Laspeyres indexes with typical element:

$$\{L_{ij}\} = \log(\mathbf{p}^i.\mathbf{q}^j/\mathbf{p}^i.\mathbf{q}^i).$$

Thus, the (logarithm) of the Laspeyres index comparing country i to country j, $\log Q^L_{ij}$, is L_{ji} and the (logarithm) of the Paasche index comparing country i to country j, $\log Q^P_{ij}$, is $-L_{ij}$. Define $\mathbf{C}$ as the matrix of logarithms of classical upper bounds to the current-weighted index, with typical element:

$$\{C_{ij}\} = \log[x^j/e_C(U^i, \mathbf{p}^j)].$$

Using this notation, the classical and fixed-weight bounds to the bilateral true indexes can be restated in logarithms as:

$$\begin{aligned} -C_{ji} \leq \log Q_{ji}^{LA} \leq L_{ij} \\ -L_{ji} \leq \log Q_{ji}^{PA} \leq C_{ij}. \end{aligned} \tag{2.10}$$

While **L** and **C** provide bounds to the (unknown) true bilateral indexes, these bounds are generally not tight. The range of indeterminacy to the bilateral true indexes, however, may be reduced via the use of third-party comparisons implemented using RP methods. Using RP methods we want to construct improved bounds to the unobservable expenditure function, and hence construct improved bounds to the unobservable bilateral welfare indexes.[16] The improved bounds to the expenditure function are denoted $e_{max}(U^i, \mathbf{p}^j)$ and $e_{min}(U^i, \mathbf{p}^j)$ where:

$$e_C(U^i, \mathbf{p}^j) \leq e_{min}(U^i, \mathbf{p}^j) \leq e(U^i, \mathbf{p}^j) \leq e_{max}(U^i, \mathbf{p}^j) \leq \mathbf{p}^j.\mathbf{q}^i.$$

Thus, $e_{max}(U^i, \mathbf{p}^j)$ is the infimum (or least upper bound) to the cost of attaining U^i when facing $\mathbf{p}^j$ (and note that $\mathbf{p}^j.\mathbf{q}^i$ is also an upper bound to this cost, but not the necessarily the least upper bound) and $e_{min}(U^i, \mathbf{p}^j)$ is the supremum (or greatest lower bound) to the cost of attaining U^i when facing $\mathbf{p}^j$ (and $e_C(U^i, \mathbf{p}^j)$ is a lower bound to this cost, but not necessarily the greatest lower bound). With the improved bounds thus determined, it is possible to construct improved bounds to the bilateral welfare indexes:

$$\begin{aligned} e_{min}(U^i, \mathbf{p}^j)/x^j \leq Q_{ij}^{LA} \leq e_{max}(U^i, \mathbf{p}^j)/x^j \\ x^i/e_{max}(U^j, \mathbf{p}^i) \leq Q_{ij}^{PA} \leq x^i/e_{min}(U^j, \mathbf{p}^i). \end{aligned} \tag{2.11}$$

Therefore, $e_{min}(U^i, \mathbf{p}^j)/x^j$ provides a *lower-bound* estimate of the base-weighted welfare index, while $e_{max}(U^i, \mathbf{p}^j)/x^j$ provides an *upper-bound* estimate of this index. Similarly, $x^i/e_{max}(U^j, \mathbf{p}^i)$ and $x^i/e_{min}(U^j, \mathbf{p}^i)$ are, respectively, lower- and upper-bound estimates of the current-weighted welfare index.

Define $\mathbf{B}^{RP}$ as the matrix of logarithms of improved upper bounds to the base-weighted welfare index (where the notation indicates that the improved bounds are constructed using RP methods), with typical element:

$$\{B_{ij}^{RP}\} = \log[e_{max}(U^j, \mathbf{p}^i)/x^i].$$

[16] Note that these improved bounds are themselves true indexes.

Similarly, define $\mathbf{C}^{RP}$ as the matrix of logarithms of upper bounds to the current-weighted welfare index, with typical element:

$$\{C_{ij}^{RP}\} = \log[x^j / e_{min}(U^i, \mathbf{p}^j)]$$

Using this notation, the bounds to the bilateral welfare indexes (2.11) can be restated (in logarithms):

$$\begin{aligned} -C_{ji}^{RP} &\leq \log Q_{ji}^{LA} \leq B_{ij}^{RP} \\ -B_{ji}^{RP} &\leq \log Q_{ji}^{PA} \leq C_{ij}^{RP}. \end{aligned} \tag{2.12}$$

In the remainder of this section, three methods for calculating $\mathbf{B}^{RP}$ and $\mathbf{C}^{RP}$ are reviewed and applied in an example.

2.4.2 GARP bounds

Varian (1982) constructs algorithms for calculating bounds to true cost-of-living and welfare indexes which are related to the GARP test of common preferences.[17] In this sub-section, a summary of Varian's algorithms for calculating the GARP bounds to true indexes is presented.

The GARP bounds are estimated using a money metric utility function (Varian calls this the *direct income compensation function*) which is defined as:

$$m(\mathbf{q}^0, \mathbf{p}) = \inf \mathbf{p}.\mathbf{q} \quad \text{subject to} \quad \mathbf{q} \text{ in } P(\mathbf{q}^0),$$

where $P(\mathbf{q}^0) = \{\mathbf{q} : u(\mathbf{q}) > u(\mathbf{q}^0)\}$. The money metric utility function can alternatively be defined as $m(\mathbf{q}^0, \mathbf{p}) = e(u(\mathbf{q}^0), \mathbf{p})$. From this latter definition, it is apparent that since the expenditure function is always increasing in utility, $m(\mathbf{q}^0, \mathbf{p})$ is a monotonic transformation of a utility function and is therefore itself a utility function.

Note that $P(\mathbf{q}^0)$ and $RP(\mathbf{q}^0)$ both describe sets of bundles which are preferred to $\mathbf{q}^0$. However, unlike $RP(\mathbf{q}^0)$, $P(\mathbf{q}^0)$ is *unobservable* since it is defined by the utility function $u(\mathbf{q})$ which is itself unobservable. However, the following is known about $P(\mathbf{q}^0)$.[18]

[17] The algorithms are implemented in a pascal program NONPAR which Professor Varian kindly made available to this author.

[18] See Section 2.2.1 for definitions of $RP(\mathbf{q}^0)$ and $RW(\mathbf{q}^0)$. The following results are all from Varian (1982).

Result 2.2. *Let $u(\mathbf{q})$ be any utility function that rationalises the data. Then for all $\mathbf{q}^0$, $RP(\mathbf{q}^0) \subset P(\mathbf{q}^0) \subset NRW(\mathbf{q}^0)$.*

Result 2.3. *Let $\mathbf{q}^0 R\mathbf{q}'$. Then $RP(\mathbf{q}^0) \subset RP(\mathbf{q}')$, $RW(\mathbf{q}^0) \supset RW(\mathbf{q}')$ and $NRW(\mathbf{q}^0) \subset NRW(\mathbf{q}')$.*

Result 2.4. *If $\mathbf{q}^i$ is in $RP(\mathbf{q}^j)$ then there exists a nonsatiated, continuous, concave, monotonic utility function that rationalises the data for which $w(\mathbf{q}^i) \geq w(\mathbf{q}^j)$. An analogous statement holds if $\mathbf{q}^i$ is in $RW(\mathbf{q}^j)$.*

Accepting that $m(\mathbf{q}^0, \mathbf{p})$ is a reasonable cardinalisation of utility, the question then is how, without using a particular parametric form of the utility or expenditure function, can it be measured? From Result 2.2 we have the best inner and outer approximations to $P(\mathbf{q}^0)$; this leads to upper and lower bounds on the money metric function:

$$m^+(\mathbf{q}^0, \mathbf{p}) = \inf \mathbf{p.q} \quad \text{subject to} \quad \mathbf{q} \text{ in } RP(\mathbf{q}^0)$$
$$m^-(\mathbf{q}^0, \mathbf{p}) = \inf \mathbf{p.q} \quad \text{subject to} \quad \mathbf{q} \text{ in } NRW(\mathbf{q}^0)$$

Result 2.5. *Let m^+ and m^- be defined as above. Then $m^-(\mathbf{q}, \mathbf{p}^0) \leq m(\mathbf{q}, \mathbf{p}^0) \leq m^+(\mathbf{q}, \mathbf{p}^0)$ for all $\mathbf{p}^0$, $\mathbf{q}$.*

Result 2.6. *Let m^+ and m^- be defined as above. Then $\mathbf{q}^i R\mathbf{q}$ implies $m^+(\mathbf{q}^i, \mathbf{p}^0) \geq m(\mathbf{q}, \mathbf{p}^0)$ and $m^-(\mathbf{q}^i, \mathbf{p}^0) \geq m^-(\mathbf{q}, \mathbf{p}^0)$.*

Result 2.5 shows that m^+ and m^- do bound the money metric function. Result 2.6 shows that m^+ and m^- are themselves utility functions that respect the RP ordering. While m^+ and m^- provide theoretically ideal bounds to the money metric function, Varian (1982) was uncertain as to whether one could easily calculate these bounds. Instead, Varian suggested two approximations to these bounds which are easily calculated via mathematical programming. The approximation to m^+ is calculated as:

$$am^+(\mathbf{q}^0, \mathbf{p}) = \inf \mathbf{p.q} \text{ subject to } \mathbf{q} \text{ is in } CM(\mathbf{q}^0),$$

where

$$CM(\mathbf{q}^0) = \text{ interior of convex hull of } \{\mathbf{q} : \mathbf{q} \geq \mathbf{q}^i, \mathbf{q}^i R\mathbf{q}^0\}.$$

Thus, to calculate $am^+(\mathbf{q}^0, \mathbf{p})$, it is necessary to calculate $\mathbf{p.q}$ only at the vertices of $\overline{CM}(\mathbf{q}^0)$ (where $\overline{CM}(\mathbf{q}^0)$ is the closure of $CM(\mathbf{q}^0)$), and find the minimum of these values. Thus the calculation of $am^+(\mathbf{q}^0, \mathbf{p})$ becomes the following minimisation problem:

$$am^+(\mathbf{q}^0, \mathbf{p}) = \min \mathbf{p}.\mathbf{q}^i \text{ subject to } \mathbf{q}^i R \mathbf{q}^0. \tag{2.13}$$

Varian (1982) proved that $CM(\mathbf{q}^0) \subseteq RP(\mathbf{q}^0)$; this implies that $am^+(\mathbf{q}^0, \mathbf{p}) \geq m^+(\mathbf{q}^0, \mathbf{p})$, and thus Varian (1982) argued that the easily calculable upper bound to the money metric $am^+(\mathbf{q}^0, \mathbf{p})$, was not as tight as the theoretically ideal bound $m^+(\mathbf{q}^0, \mathbf{p})$. However, Knoblauch (1992) proved that $RP(\mathbf{q}^0) \subseteq \overline{CM}(\mathbf{q}^0)$, so that $m^+(\mathbf{q}^0, \mathbf{p}) = am^+(\mathbf{q}^0, \mathbf{p})$ and therefore the theoretically ideal upper bound to the money metric function is in fact easily calculable using (2.13).

Using the same logic, it is apparent that $m^-(\mathbf{q}^0, \mathbf{p})$ is calculated as the following minimisation problem:

$$m^-(\mathbf{q}^0, \mathbf{p}) = \min \mathbf{p}.\mathbf{q}^j \text{ subject to } \mathbf{p}^i.\mathbf{q}^j > \mathbf{p}^i.\mathbf{q}^i \text{ for all } \mathbf{q}^i \text{ such that } \mathbf{q}^0 R \mathbf{q}^i. \tag{2.14}$$

The GARP approximations to the expenditure function defined in (2.13) and (2.14) can be used to construct improved bounds to the bilateral true welfare indexes. Define $\mathbf{B}^G$ as the matrix of logarithms of GARP upper bounds to the base-weighted true index, with typical element:

$$\{B^G_{ij}\} = \log[m^+(\mathbf{q}^j, \mathbf{p}^i)/x^i].$$

Similarly, define $\mathbf{C}^G$ as the matrix of logarithms of GARP upper bounds to the current-weighted true index, with typical element:

$$\{C^G_{ij}\} = \log[x^j/m^-(\mathbf{q}^i, \mathbf{p}^j)].$$

Example

For the 6 country data, the fixed-weight and classical bounds to the bilateral true welfare indexes are contained in the following matrices:

$$\mathbf{L} = \begin{bmatrix} 0.000 & 1.365 & 0.511 & 0.693 & 1.126 & 0.981 \\ -0.736 & 0.000 & -0.470 & 0.318 & -0.539 & -0.234 \\ -0.357 & 0.642 & 0.000 & 0.588 & 0.262 & 0.336 \\ -0.568 & 0.916 & 0.000 & 0.000 & 0.710 & 0.511 \\ 0.248 & 0.754 & 0.446 & 1.371 & 0.000 & 0.595 \\ -0.405 & 0.189 & -0.182 & 0.693 & -0.470 & 0.000 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 0.000 & 1.792 & 0.916 & 0.916 & 1.386 & 1.792 \\ 0.000 & 0.000 & -0.182 & 0.916 & -0.405 & 0.000 \\ 0.182 & 0.875 & 0.000 & 1.099 & 0.470 & 0.875 \\ -0.405 & 1.386 & 0.511 & 0.000 & 0.981 & 1.386 \\ 1.099 & 1.099 & 0.916 & 2.015 & 0.000 & 0.693 \\ 0.405 & 0.405 & 0.223 & 1.322 & -0.405 & 0.000 \end{bmatrix}$$

The Laspeyres index provides the upper bound to the base-weighted welfare index and thus, from **L**, the welfare of country 5 is at most $e^{0.262} = 1.30$ times the welfare of country 3, when using $\mathbf{p}^3$ in the welfare measurement. The Paasche index provides the lower bound to the current-weighted welfare index and hence the welfare of country 5 is at least $e^{-0.446} = 0.64$ times the welfare of country 3, when using $\mathbf{p}^5$ in the welfare measurement.

The bounds to the base-weighted true index, with country 3 as base are therefore:

$$\begin{aligned} -C_{13} = \; & -0.916 \leq \log Q_{21}^{LA} \leq -0.357 = B_{31} \\ -C_{23} = \; & 0.182 \leq \log Q_{21}^{LA} \leq 0.642 = B_{32} \\ -C_{43} = \; & -0.511 \leq \log Q_{31}^{LA} \leq 0.588 = B_{34} \\ -C_{53} = \; & -0.916 \leq \log Q_{41}^{LA} \leq 0.262 = B_{35} \\ -C_{63} = \; & -0.223 \leq \log Q_{51}^{LA} \leq -0.336 = B_{36}. \end{aligned}$$

while the bounds to the current-weighted true index are:

$$\begin{aligned} -B_{13} = \; & -0.511 \leq \log Q_{21}^{LA} \leq 0.182 = C_{31} \\ -B_{23} = \; & 0.470 \leq \log Q_{21}^{LA} \leq 0.875 = C_{32} \\ -B_{43} = \; & 0.000 \leq \log Q_{31}^{LA} \leq 1.099 = C_{34} \\ -B_{53} = \; & -0.446 \leq \log Q_{41}^{LA} \leq 0.470 = C_{35} \\ -B_{63} = \; & 0.182 \leq \log Q_{51}^{LA} \leq 0.875 = C_{36}. \end{aligned}$$

Varian's (1982) GARP bounds were then calculated for the 6 country data:

$$\mathbf{B}^G = \begin{bmatrix} 0.000 & 1.365 & 0.511 & 0.693 & 0.981 & 0.981 \\ -0.736 & 0.000 & -0.470 & 0.318 & -0.539 & -0.234 \\ -0.357 & 0.642 & 0.000 & 0.588 & 0.262 & 0.336 \\ -0.568 & 0.916 & 0.000 & 0.000 & 0.511 & 0.511 \\ 0.248 & 0.754 & 0.446 & 1.371 & 0.000 & 0.595 \\ -0.405 & 0.189 & -0.182 & 0.693 & -0.470 & 0.000 \end{bmatrix}$$

$$\mathbf{C}^G = \begin{bmatrix} 0.000 & 1.792 & 0.916 & 0.916 & 1.386 & 1.792 \\ 0.000 & 0.000 & -0.182 & 0.916 & -0.405 & 0.000 \\ 0.000 & 0.875 & 0.000 & 0.916 & 0.470 & 0.875 \\ -0.511 & 0.875 & 0.000 & 0.000 & 0.470 & 0.875 \\ 1.099 & 1.099 & 0.916 & 2.015 & 0.000 & 0.693 \\ 0.000 & 0.345 & 0.000 & 0.916 & -0.405 & 0.000 \end{bmatrix}$$

Several of the bounds have been improved as a result of using the GARP relations. For example, the upper bound to $\log Q_{51}^{LA}$ has been tightened from $L_{15} = 1.126$ to $B_{15}^G = 0.981$ and the upper bound to $\log Q_{24}^{PA}$ has been tightened from $C_{42} = 1.386$ to $C_{42}^G = 0.875$.

2.4.3 Improved GARP bounds

For those countries which satisfy the improved GARP test of common preferences of Blundell, Browning, and Crawford (1998), it is possible to construct improved (i.e. tighter) bounds to the bilateral true welfare indexes. Essentially, the RP and RW algorithms of Blundell, Browning, and Crawford (1998) use expansion path information to determine the regions $RP(\mathbf{q})$ and $NRW(\mathbf{q})$. Then, using the results of Varian (1982), it is simply a matter of calculating the value of the bundles on the vertices of $RP(\mathbf{q})$ and $NRW(\mathbf{q})$ - these values will be the approximations to the expenditure function, which are then used in computing the bilateral bounds.

The improved GARP bounds are therefore calculated using the following approximations to the expenditure function:

$$\begin{aligned} m^{i+}(\mathbf{q}^0, \mathbf{p}) &= \min \mathbf{p}.\mathbf{q} \quad \text{subject to} \quad \mathbf{q} \text{ in } RP(\mathbf{q}^0) \\ m^{i-}(\mathbf{q}^0, \mathbf{p}) &= \min \mathbf{p}.\mathbf{q} \quad \text{subject to} \quad \mathbf{q} \text{ in } NRW(\mathbf{q}^0). \end{aligned}$$

where $RP(\mathbf{q}^0)$ and $NRW(\mathbf{q}^0)$ are determined using the BBC algorithms. Define $\mathbf{B}^{IG}$ as the matrix of logarithms of improved GARP upper bounds to the base-weighted welfare index, with typical element:

$$\{B_{ij}^{IG}\} = \log[m^{I+}(\mathbf{q}^j, \mathbf{p}^i)/x^i].$$

Similarly, define $\mathbf{C}^{IG}$ as the matrix of logarithms of improved GARP upper bounds to the current-weighted welfare index, with typical element:

$$\{C_{ij}^{IG}\} = \log[x^j / m^{I-}(\mathbf{q}^i, \mathbf{p}^j)].$$

Example

It was shown above that the improved GARP test of common preferences is satisfied when country 6 is omitted from the example data set. The fact that the improved test of GARP is satisfied implies that it is possible to construct improved inner and outer bounds on in the indifference curves for all countries in the data set. The *RP* and *RW* bounds to the indifference curve for country 3 are shown in Figure 2.17.

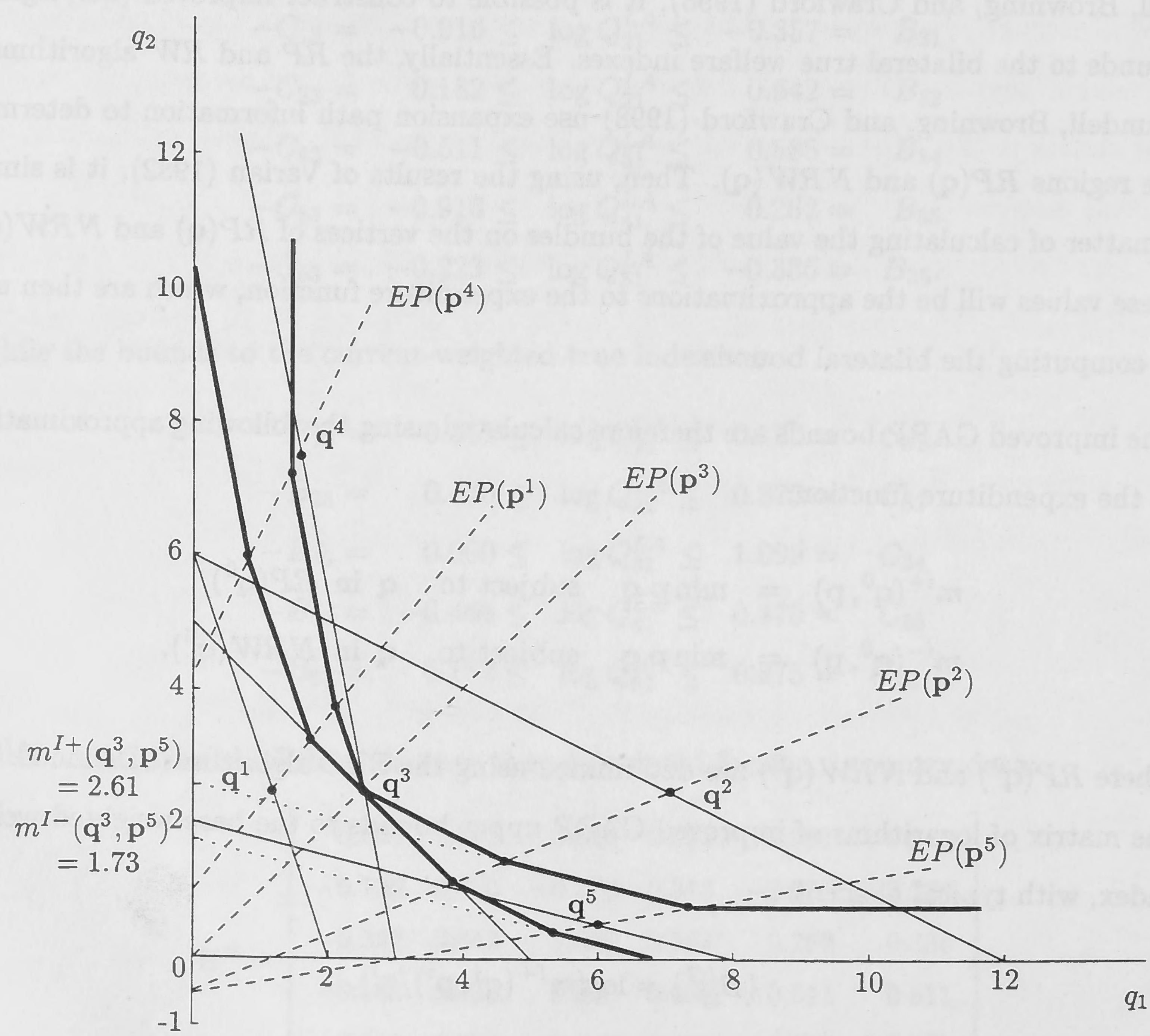

Figure 2.17: The *RP* and *RW* bounds for country 3

The matrix of log Laspeyres indexes for countries 1-5 is:

$$\mathbf{L} = \begin{bmatrix} 0.000 & 1.365 & 0.511 & 0.693 & 1.126 \\ -0.736 & 0.000 & -0.470 & 0.318 & -0.539 \\ -0.357 & 0.642 & 0.000 & 0.588 & 0.262 \\ -0.568 & 0.916 & 0.000 & 0.000 & 0.710 \\ 0.248 & 0.754 & 0.446 & 1.371 & 0.000 \end{bmatrix}$$

The matrix of classical upper bounds to the current-weighted welfare index is:

$$\mathbf{C} = \begin{bmatrix} 0.000 & 1.792 & 0.916 & 0.916 & 1.386 \\ 0.000 & 0.000 & -0.182 & 0.916 & -0.405 \\ 0.182 & 0.875 & 0.000 & 1.099 & 0.470 \\ -0.405 & 1.386 & 0.511 & 0.000 & 0.981 \\ 1.099 & 1.099 & 0.916 & 2.015 & 0.000 \end{bmatrix}$$

Varian's (1982) GARP bounds for the 5 country data are:

$$\mathbf{B}^G = \begin{bmatrix} 0.000 & 1.365 & 0.511 & 0.693 & 1.126 \\ -0.736 & 0.000 & -0.470 & 0.318 & -0.539 \\ -0.357 & 0.642 & 0.000 & 0.588 & 0.262 \\ -0.568 & 0.916 & 0.000 & 0.000 & 0.710 \\ 0.248 & 0.754 & 0.446 & 1.371 & 0.000 \end{bmatrix}$$

$$\mathbf{C}^G = \begin{bmatrix} 0.000 & 1.792 & 0.916 & 0.916 & 1.386 \\ 0.000 & 0.000 & -0.182 & 0.916 & -0.405 \\ 0.000 & 0.875 & 0.000 & 0.916 & 0.470 \\ -0.511 & 0.875 & 0.000 & 0.000 & 0.470 \\ 1.099 & 1.099 & 0.916 & 2.015 & 0.000 \end{bmatrix}$$

The improved GARP bounds of Blundell, Browning, and Crawford (1998) for the 5 country data are:

$$\mathbf{B}^{IG} = \begin{bmatrix} 0.000 & 1.153 & 0.511 & 0.693 & 0.634 \\ -0.827 & 0.000 & -0.470 & -0.208 & -0.539 \\ -0.357 & 0.642 & 0.000 & 0.262 & 0.123 \\ -0.568 & 0.624 & -0.034 & 0.000 & 0.093 \\ -0.114 & 0.754 & 0.265 & 0.539 & 0.000 \end{bmatrix}$$

$$C^{IG} = \begin{bmatrix} 0.000 & 1.222 & 0.511 & 0.804 & 0.783 \\ -0.917 & 0.000 & -0.470 & -0.245 & -0.578 \\ -0.407 & 0.670 & 0.000 & 0.323 & 0.146 \\ -0.543 & 0.635 & -0.033 & 0.000 & 0.111 \\ -0.124 & 0.719 & 0.249 & 0.653 & 0.000 \end{bmatrix}$$

The incorporation of expansion path information has led to the bounds to the bilateral welfare indexes being significantly tightened. Several of the bounds have been improved as a result of using the GARP relations. For example, the GARP upper bound to $\log Q_{51}^{LA}$ is $B_{15}^{G} = 1.126$ (which is equivalent to the fixed-weight bound), while the improved GARP upper bound is $B_{15}^{IG} = 0.634$. Similarly, the GARP upper bound to $\log Q_{24}^{PA}$ is $C_{42}^{G} = 0.875$, while the improved GARP upper bound is $C_{42}^{IG} = 0.635$.

2.4.4 Afriat envelope bounds

As noted above, the Afriat numbers calculated using the mathematical program (2.9) are not unique; underlying this is the fact that for any set of demand data there will be a whole family of utility functions that rationalise the data. However, Afriat (1987) shows that the family of utility functions may be bounded by two particular representations of the utility function, which are known as the *inner* and *outer envelope functions* for the class of compatible utility functions associated with any particular solution to the Afriat inequalities. In this subsection the Afriat envelope approach to making multilateral approximations to bilateral true welfare indexes is outlined.[19]

Conditional inner and outer envelope functions

For given sets $\boldsymbol{W}$ and $\boldsymbol{\lambda}$ satisfying (2.1) and (2.2), the inner envelope function is a monotonic concave polytope function and is defined:

$$(2.15) \qquad w_I(\mathbf{q}, \boldsymbol{W}) = \max_{\boldsymbol{\theta}} \Big[\sum_{i \in N} W^i \theta^i : \ \sum_{i \in N} \mathbf{q}^i \theta^i \leq \mathbf{q}, \ \sum_{i \in N} \theta^i = 1, \ \theta^i \geq 0, \ i \in N \Big].$$

[19]The following presentation is based on Chavas and Cox (1997) who use Afriat envelope functions to construct cost-of-living indexes using aggregate U.S. consumption data.

The outer envelope function is a monotonic concave polyhedral function[20] and is defined:

$$w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda}) = \min_i [W^i + \lambda^i(\mathbf{v}^i.\mathbf{q} - 1) : \ i \in N]. \tag{2.16}$$

Note that (2.16) can alternatively be expressed as the following primal linear programming problem:

$$w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda}) = \max_{\gamma} [\gamma : \ \gamma \leq W^i + \lambda^i(\mathbf{v}^i.\mathbf{q} - 1), \ i \in N]. \tag{2.17}$$

Since by construction $w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda})$ and $w_I(\mathbf{q}, \boldsymbol{W})$ satisfy the Afriat inequalities, the following Lemma applies.

Lemma 2.1. *$w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda})$ and $w_I(\mathbf{q}, \boldsymbol{W})$ are representations of consumer preferences which rationalise the demand data.*

Afriat (1987) has shown that the functions $w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda})$ and $w_I(\mathbf{q}, \boldsymbol{W})$ provide bounds on the family of utility functions that are compatible with the demand data.

Result 2.7. Afriat (1987) *For $\boldsymbol{W}$ and $\boldsymbol{\lambda}$ satisfying* (2.1) *and* (2.2), *the functions $w_I(\mathbf{q}, \boldsymbol{W})$ and $w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda})$ provide, respectively, the outer bound and inner bound representation of consumer preferences:*

$$w_I(\mathbf{q}, \boldsymbol{W}) \leq w(\mathbf{q}) \leq w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda}),$$

where $w(\mathbf{q})$ is any concave, monotonic, continuous, non-satiated utility function that rationalises the data and satisfying $w(\mathbf{q}^i) = W^i$ for all $i \in N$

Result 2.7 shows that $w_I(\mathbf{q}, \boldsymbol{W})$ and $w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda})$ provide the tightest possible bounds on all possible concave and non-satiated utility representations of $u(\mathbf{q})$. However, it should be emphasised that these bounds are *conditional* on particular $\boldsymbol{W}$ and $\boldsymbol{\lambda}$ satisfying the Afriat inequalities (2.1) and (2.2).

Unconditional inner and outer envelope functions

As previously mentioned, in general there will be more than one solution to the Afriat inequalities. This means that the conditional inner and outer envelope functions presented

[20]The function (2.16) is a concave polyhedral because it is not defined for all prices; it is only defined for the finite number of price vectors in the data set.

above will not be unique. Chavas and Cox (1997) therefore suggest the need to calculate *unconditional* bounds for $w(\mathbf{q})$, i.e. bounds that do not depend on the values taken by (W, λ).

Note that the concave function $w(\mathbf{q})$ is only defined up to a positive linear transformation and without loss of generality, the Afriat inequalities (2.1) and (2.2) can alternatively be written as:

$$\lambda^i \geq \epsilon, \quad i \in N \tag{2.18}$$

$$W^i \leq W^j + \lambda^j(\mathbf{v}^j.\mathbf{q}^i - 1), \quad i, j \in N \tag{2.19}$$

$$W^b = 0 \tag{2.20}$$

$$-m \leq W^i \leq m, \quad i \in N. \tag{2.21}$$

where ϵ is a small positive scalar, W^b is the utility level for some base observation $b \in N$, and m is some positive, finite scalar. Equations (2.18) and (2.19) correspond to (2.1) and (2.2). Equation (2.20) is a normalisation which arbitrarily fixes the utility level for some base observation b and (2.21) is used to ensure that the utility levels W^i remain bounded. It is apparent that (2.18), (2.19), (2.20), (2.21) are therefore an equivalent means of imposing the Afriat inequalities (2.1) and (2.2).

We are now able to define *unconditional* bounds at point $\mathbf{q}$:

$$w_{UI}(\mathbf{q}) = \min_{\boldsymbol{W}}[w_I(\mathbf{q}, \boldsymbol{W}) : \text{ Eq. } (2.18) - (2.21)]. \tag{2.22}$$

and

$$w_{UO}(\mathbf{q}) = \max_{\boldsymbol{W},\boldsymbol{\lambda}}[w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda}) : \text{ Eq. } (2.18) - (2.21)]. \tag{2.23}$$

Result 2.8. Chavas and Cox (1997) *Given* (2.18)-(2.21), *the functions* w_{UI} *in* (2.22) *and* w_{UO} *in* (2.23) *provide, respectively, the unconditional inner and outer bound representations of consumer preferences at point* $\mathbf{q}$:

$$w_{UI}(\mathbf{q}) \leq w(\mathbf{q}) \leq w_{UO}(\mathbf{q}).$$

Result 2.8 shows that the two functions $w_{UI}(\mathbf{q})$ and $w_{UO}(\mathbf{q})$ provide the widest possible bounds on all possible concave and non-satiated utility representations of $u(\mathbf{q})$.[21]

[21] Note, however, that these bounds have been constructed using RP relations and hence they will generally be tighter than the fixed-weight and classical bounds.

Inner and outer expenditure functions

The two unconditional bounds $w_{UI}(\mathbf{q})$ and $w_{UO}(\mathbf{q})$ provide a basis for conducting welfare or standard of living comparisons via the construction of nonparametric bounds to the expenditure function. First, note that $w_{UI}(\mathbf{q}) = w_I(\mathbf{q}, \boldsymbol{W}_I)$, where $\boldsymbol{W}_I = \{W_I^1, \ldots, W_I^N\}$ is the unconditional inner-bound utility representation obtained in (2.22). Thus, using (2.15), the unconditional inner-bound representation of the expenditure function is:

$$\begin{aligned} e_I(W^s, \mathbf{p}) &= \min_{\mathbf{q}}[\mathbf{p}.\mathbf{q} : \; w_I(\mathbf{q}, \boldsymbol{W}_I) \geq W^s; \; \mathbf{q} \geq 0] \\ &= \min_{\mathbf{q},\boldsymbol{\theta}}[\mathbf{p}.\mathbf{q} : \sum_{i\in N} W_I^i \theta^i \geq W^s; \; \sum_{i \in N} \mathbf{q}^i \theta^i \leq \mathbf{q}; \\ &\qquad \sum_{i\in N} \theta^i = 1; \; \theta^i \geq 0; \; i \in N; \; \mathbf{q} \geq 0], \end{aligned} \tag{2.24}$$

where $W^s = W_I^s$ obtained in (2.22). Equation (2.24) is a linear program which can be used to find the expenditure function $e_I(W^s, \mathbf{p})$ associated with $w_{UI}(\mathbf{q})$, the unconditional outer-bound representation of preferences, when prices are $\mathbf{p}$.

Similarly, note that $w_{UO}(\mathbf{q}) = w_O(\mathbf{q}, \boldsymbol{W}_O, \boldsymbol{\lambda}_O)$, where $\boldsymbol{W}_O = \{W_O^1, \ldots, W_O^N\}$ and $\boldsymbol{\lambda}_O = \{\lambda_O^1, \ldots, \lambda_O^N\}$ are the outer-bound representations obtained in (2.23). Thus, using (2.17), the unconditional outer-bound representation of the expenditure function is:

$$\begin{aligned} e_O(W^s, \mathbf{p}) &= \min_{\mathbf{q}}[\mathbf{p}.\mathbf{q} : \; w_O(\mathbf{q}, \boldsymbol{W}_O, \boldsymbol{\lambda}_O) \geq W^s; \; \mathbf{q} \geq 0] \\ &= \min_{\mathbf{q},\gamma}[\mathbf{p}.\mathbf{q} : \; \gamma \geq W^s; \; \gamma \leq W_O^i + \lambda_O^i(\mathbf{v}^i.\mathbf{q} - 1); \; i \in N; \; \mathbf{q} \geq 0] \\ &= \min_{\mathbf{q}}[\mathbf{p}.\mathbf{q} : \; W^s \leq W_O^i + \lambda_O^i(\mathbf{v}^i.\mathbf{q} - 1); \; i \in N; \; \mathbf{q} \geq 0], \end{aligned} \tag{2.25}$$

where $W^s = W_O^s$ is obtained in (2.23). Equation (2.25) is a linear program which can be used to find the expenditure function $e_O(W^s, \mathbf{p})$ associated with $w_{UO}(\mathbf{q})$, the unconditional outer-bound representation of preferences, when prices are $\mathbf{p}$.

The expenditure functions $e_I(W^s, \mathbf{p})$ in (2.24) and $e_O(W^s, \mathbf{p})$ in (2.25) provide nonparametric bounds to the (unknown) true expenditure function:

Result 2.9. CHAVAS AND COX (1997) *Let $\{W_I^s, \; s \in N\}$ be the solution of* (2.22) *and $\{W_O^s, \; s \in N\}$ be the solution of* (2.23). *Then the functions $e_I(W^s, \mathbf{p})$ in* (2.24) *and $e_O(W^s, \mathbf{p})$ in* (2.25) *provide, respectively, the inner-bound and outer-bound representation of the expenditure function:*

$$e_O(W^s, \mathbf{p}) \leq e(W^s, \mathbf{p}) \leq e_I(W^s, \mathbf{p}), \tag{2.26}$$

where $e(W^s, \mathbf{p})$ *is an expenditure function that rationalises the data and is defined as* $e(W^s, \mathbf{p}) \equiv \min_{\mathbf{q}} \mathbf{p}.\mathbf{q}$ *s.t.* $w(\mathbf{q}) \geq W^s$; $\mathbf{q} \geq 0$, *and* $w(\mathbf{q})$ *is a concave, non-satiated utility function which rationalises the data.*

By evaluating the expenditure functions $e_I(W^s, \mathbf{p})$ and $e_O(W^s, \mathbf{p})$ for different price vectors, one can construct nonparametric bounds to welfare indexes. From (2.26), $e_I(W^i, \mathbf{p}^j)$ is the highest possible cost of attaining utility W^i when prices are $\mathbf{p}^j$, and $e_O(W^i, \mathbf{p}^j)$ is the lowest possible cost of attaining utility W^i when prices are $\mathbf{p}^j$. Thus, the unconditional inner and outer bounds to the expenditure function can be used to provide bounds to the welfare indexes:

$$
\begin{aligned}
& e_O(W^i, \mathbf{p}^j)/x^j \leq Q_{ij}^{LA} \leq e_I(W^i, \mathbf{p}^j)/x^j \\
& x^i/e_I(W^j, \mathbf{p}^i) \leq Q_{ij}^{PA} \leq x^i/e_O(W^j, \mathbf{p}^i).
\end{aligned} \tag{2.27}
$$

Therefore, $e_O(W^i, \mathbf{p}^j)/x^j$ provides a *lower-bound* estimated of the base-weighted welfare index, while $e_I(W^i, \mathbf{p}^j)/x^j$ provides an *upper-bound* estimated of this index. Similarly, $x^i/e_I(W^j, \mathbf{p}^i)$ and $x^i/e_O(W^j, \mathbf{p}^i)$ are, respectively, lower- and upper-bound estimates of the current-weighted welfare index.

Now define $\mathbf{B}^A$ as the matrix of logarithms of (Afriat envelope) improved upper bounds to the base-weighted welfare index, with typical element:

$$\{B_{ij}^A\} = \log[e_I(W^j, \mathbf{p}^i)/x^i].$$

Similarly, define $\mathbf{C}^A$ as the matrix of logarithms of (Afriat envelope) improved upper bounds to the current-weighted welfare index, with typical element:

$$\{C_{ij}^A\} = \log[x^j/e_O(W^i, \mathbf{p}^j)].$$

Using this notation, the bounds to the bilateral welfare indexes (2.27) can be restated (in logarithms):

$$
\begin{aligned}
& -C_{ji}^A \leq \log Q_{ji}^{LA} \leq B_{ij}^A \\
& -B_{ji}^A \leq \log Q_{ji}^{PA} \leq C_{ij}^A.
\end{aligned} \tag{2.28}
$$

Example

The Afriat envelope functions and the implied bounds to the bilateral welfare indexes are now constructed for the 6 country data. As shown in Section 2.2.4, when the linear program (2.9) was run with country 3 as the base country it attained the lower bound of zero (and thus the data can be rationalised by a utility function). The calculated Afriat numbers are: $W^1 = -0.663$, $W^2 = 0.750$, $W^3 = 0$, $W^4 = 0.341$, $W^5 = -0.268$ $W^6 = 0.431$; $\lambda^1 = 1.005$, $\lambda^2 = 1.362$, $\lambda^3 = 1.078$, $\lambda^4 = 1.717$, $\lambda^5 = 1.445$ and $\lambda^6 = 1.529$.

Conditional inner and outer envelope functions The inner envelope function conditional on these Afriat numbers are shown in Figures (2.18) and (2.19), while the conditional outer envelope function is shown in Figures (2.20) and (2.21). An overhead view, which shows the conditional inner- and outer-bounds to the indifference curve for country 3 is shown in Figure 2.22.

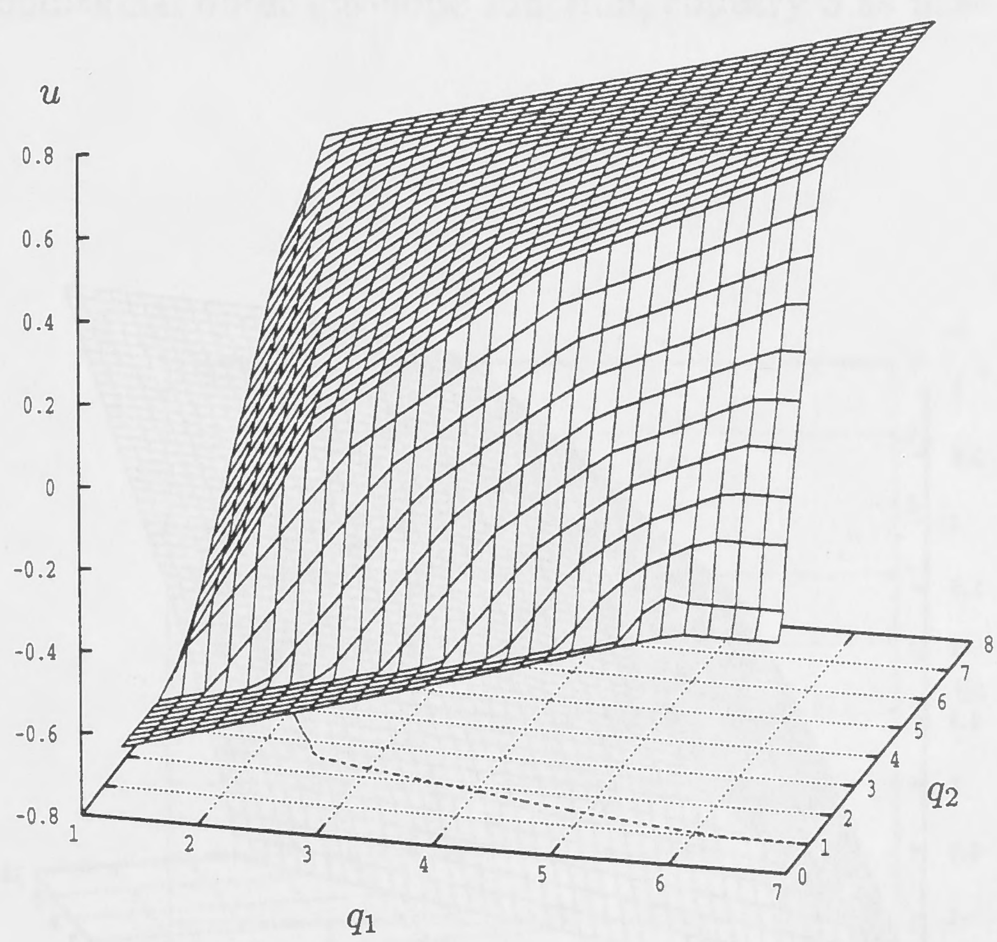

Figure 2.18: Conditional inner envelope function, country 3 as base

Unconditional inner and outer envelope functions While the indifference curves in Figure 2.22 are constructed from the Afriat numbers and thus are consistent with

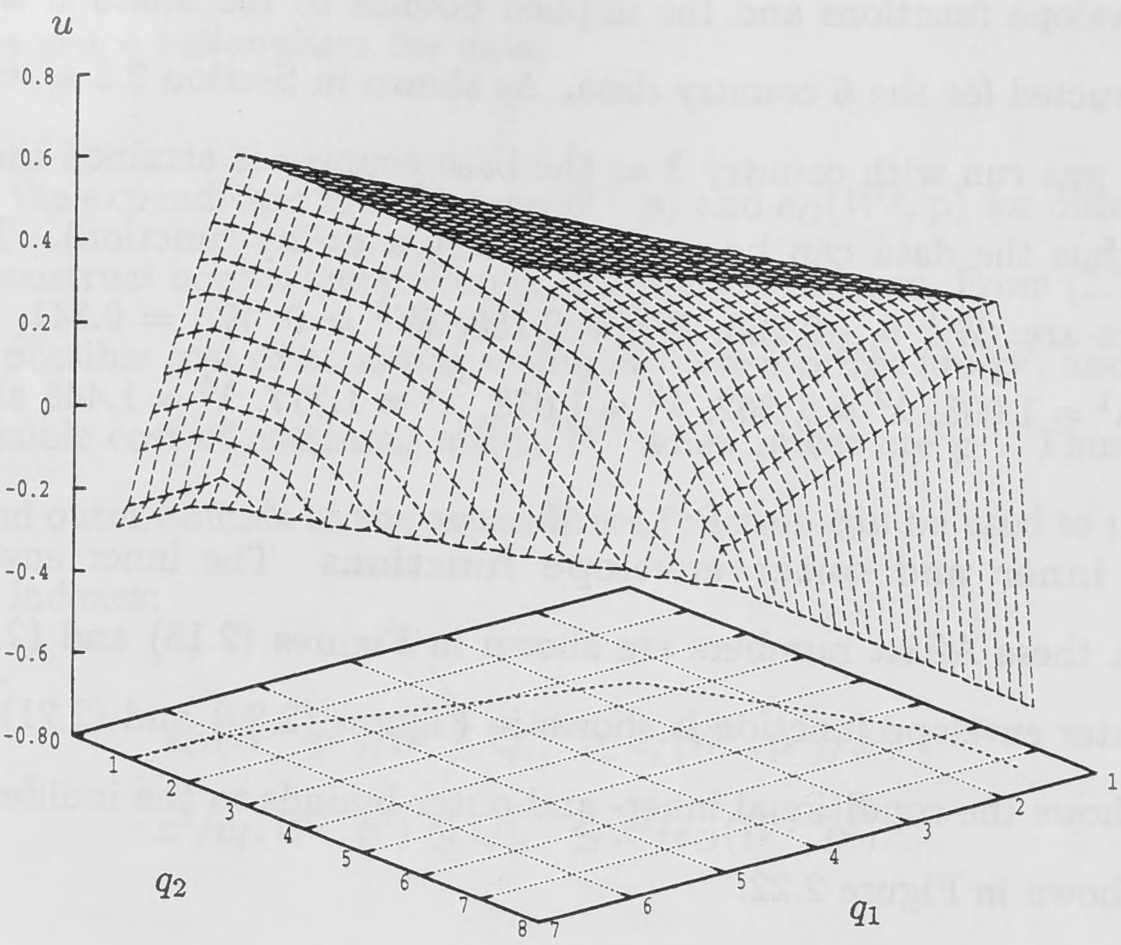

Figure 2.19: Conditional inner envelope function, country 3 as base (alternative view)

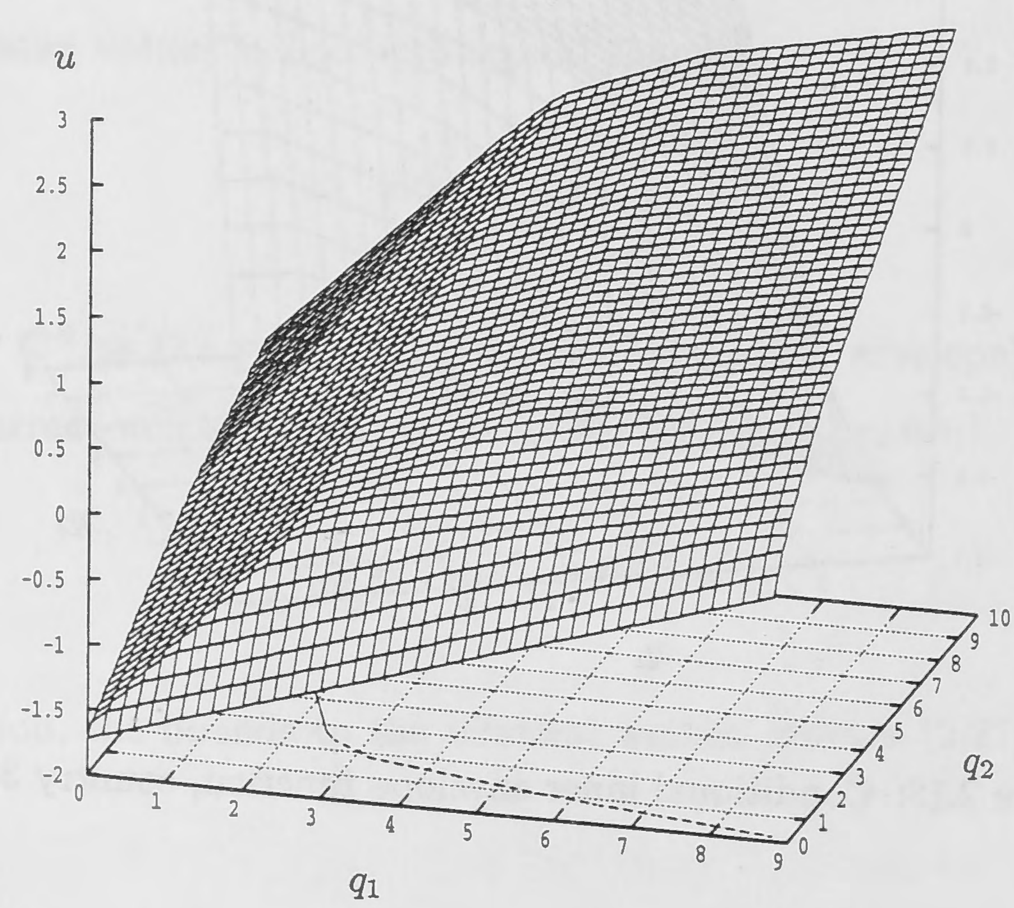

Figure 2.20: Conditional outer envelope function, country 3 as base

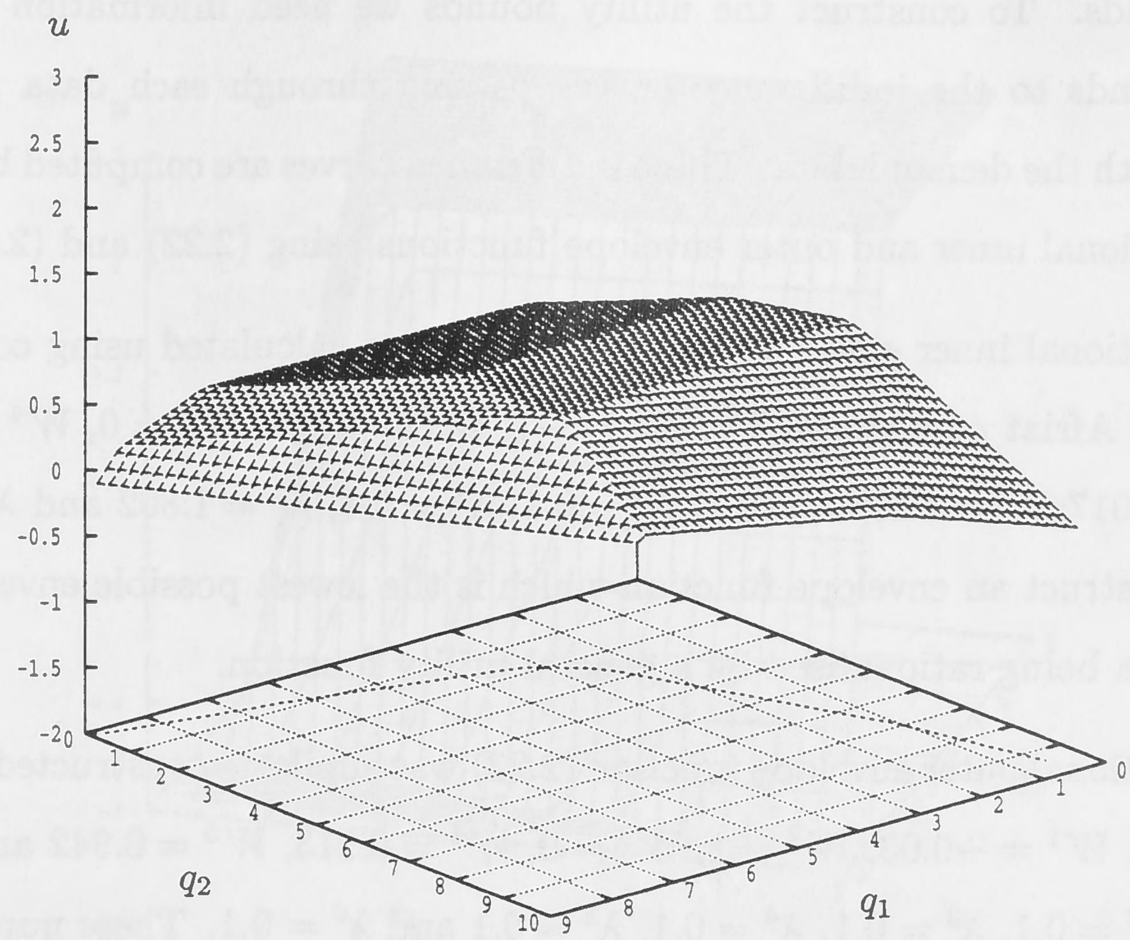

Figure 2.21: Conditional outer envelope function, country 3 as base (alternative view)

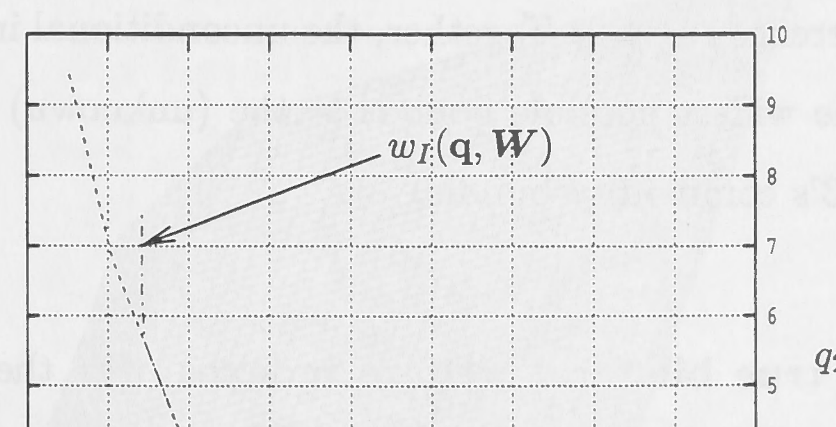

Figure 2.22: Conditional inner and outer indifference curves for country 3

a utility function which rationalises the data, the fact that they are conditional on a particular set of Afriat numbers implies they are of limited use in the construction of welfare bounds. To construct the utility bounds we need information on the widest possible bounds to the indifference curves passing through each data point that are consistent with the demand data. These indifference curves are computed by constructing the unconditional inner and outer envelope functions using (2.22) and (2.23).

The unconditional inner envelope function (2.22) was calculated using country 3 as the base and the Afriat numbers are:[22] $W^1 = -1$, $W^2 = 0.038$, $W^3 = 0$, $W^4 = 0$, $W^5 = -1$ and $W^6 = 0.017$; $\lambda^1 = 1.5$, $\lambda^2 = 0.1$, $\lambda^3 = 0.1$, $\lambda^4 = 0.1$, $\lambda^5 = 1.802$ and $\lambda^6 = 0.1$. These numbers construct an envelope function which is the lowest possible envelope consistent with the data being rationalised by a general utility function.

The unconditional outer envelope function (2.23) was similarly constructed and the Afriat numbers are: $W^1 = -0.03$, $W^2 = 1$, $W^3 = 0$, $W^4 = 0.913$, $W^5 = 0.942$ and $W^6 = 0.979$; $\lambda^1 = 0.11$, $\lambda^2 = 0.1$, $\lambda^3 = 0.1$, $\lambda^4 = 0.1$, $\lambda^5 = 0.1$ and $\lambda^6 = 0.1$. These numbers construct an envelope function which is the highest possible envelope consistent with the data being rationalised by a general utility function. The unconditional inner and outer envelope functions are presented in Figures (2.23) and (2.24).[23]

The unconditional inner and outer indifference curves passing through country 3's commodity bundle are shown in Figure 2.25 (which, for comparative purposes, also shows the conditional indifference curves). Together, the unconditional inner and outer indifference curves provide the widest possible bounds to the (unknown) indifference curve passing through country 3's commodity bundle.

Bounds to the true bilateral welfare indexes For the 6 country example data, the fixed-weight and classical bounds (**L** and **C**) were shown above, as were the GARP bounds ($\mathbf{B}^G$ and $\mathbf{C}^G$). The Afriat envelope bounds to the bilateral welfare indexes are now constructed for the 6 country data using the unconditional inner-bound and outer-bound representations of the expenditure function, (2.24) and (2.25). Given that the data can be rationalised by a utility function, it was shown in (2.28) that the tightest

[22] A value of $m = 1$ was used in the constraint (2.21).

[23] Note that the unconditional inner envelope function is the same as the conditional inner envelope function, conditional on the unconditional (inner) Afriat numbers, and the unconditional outer envelope function is the same as the conditional outer envelope function, conditional on the unconditional (outer) Afriat numbers.

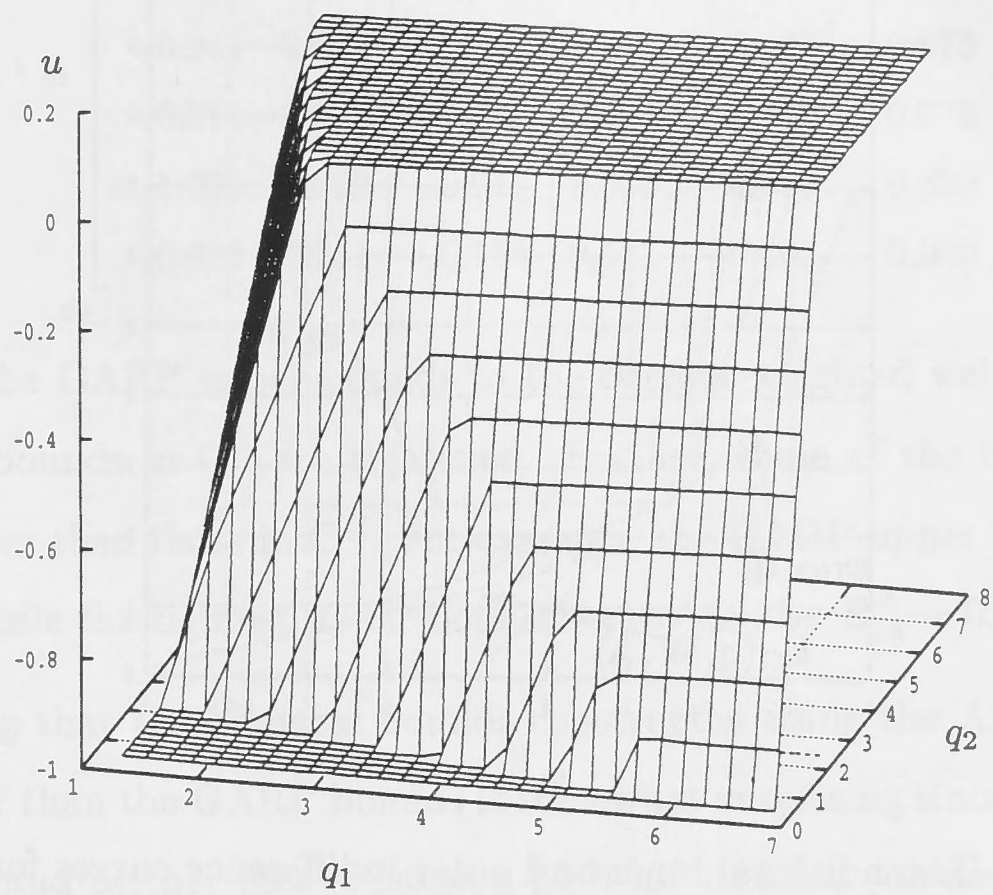

Figure 2.23: Unconditional inner envelope function, country 3 as base

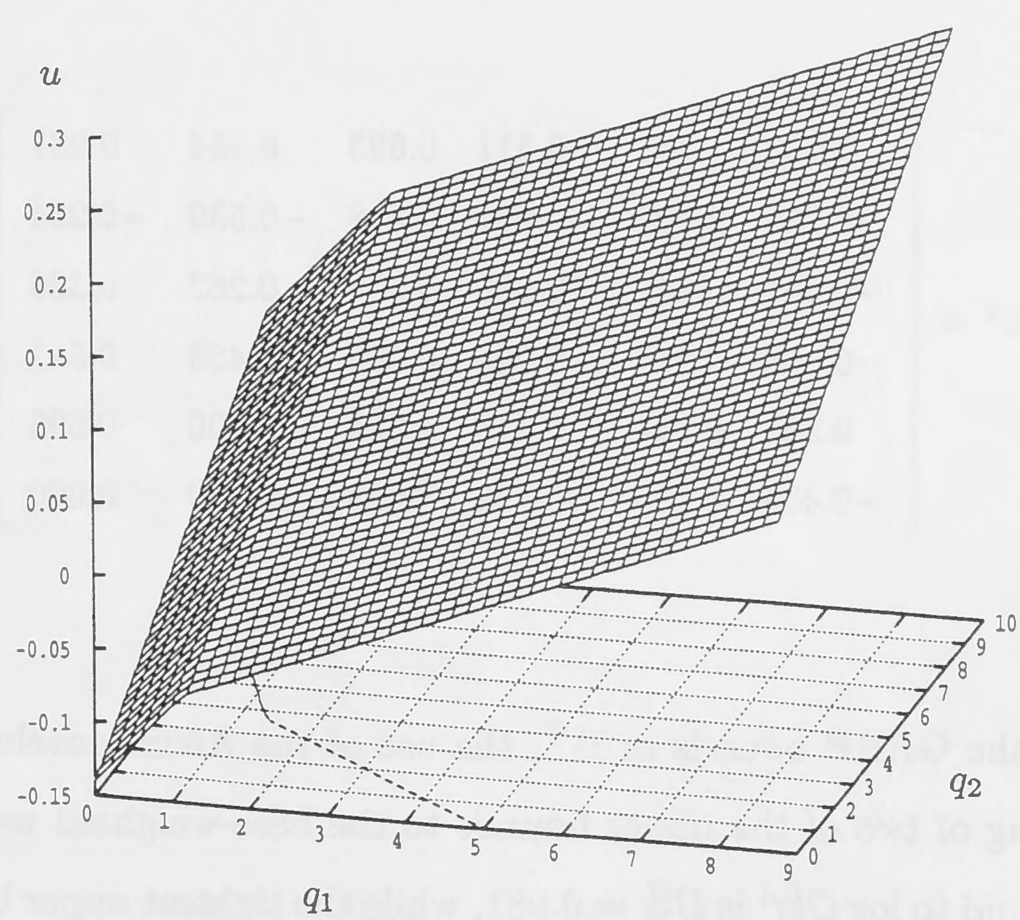

Figure 2.24: Unconditional outer envelope function, country 3 as base

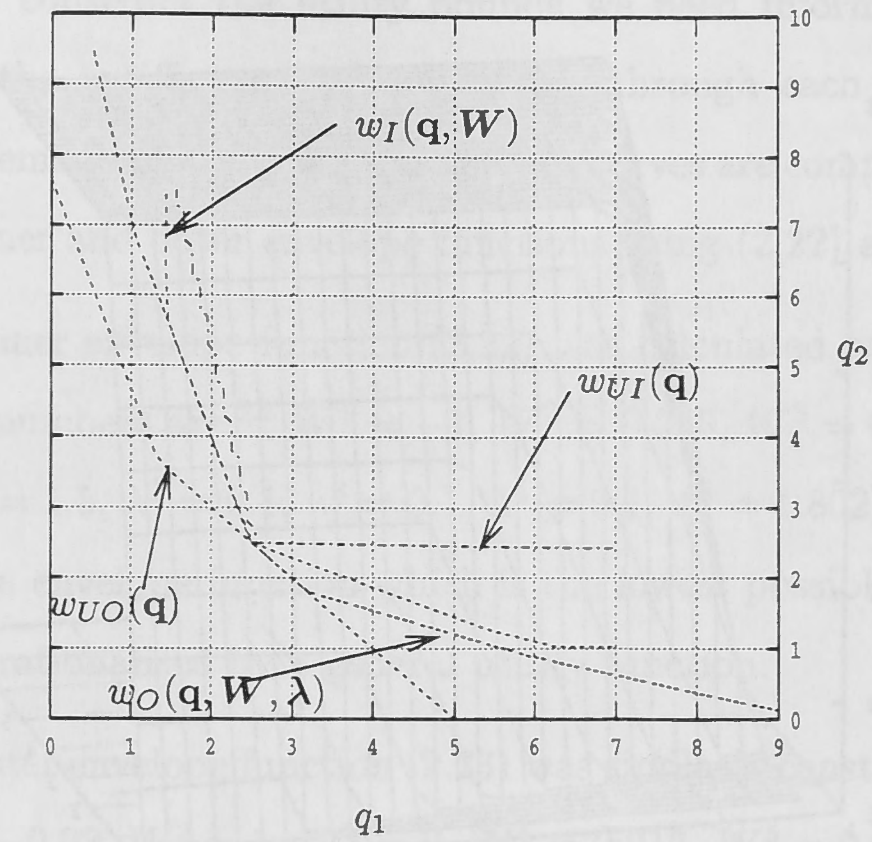

Figure 2.25: Unconditional inner and outer indifference curves for country 3

possible upper bounds to the base-weighted welfare index are provided by the elements of $\mathbf{B}^A$, which for this example is:

$$\mathbf{B}^A = \begin{bmatrix} 0.000 & 1.365 & 0.511 & 0.693 & 0.954 & 0.981 \\ -0.736 & 0.000 & -0.470 & 0.318 & -0.539 & -0.234 \\ -0.357 & 0.642 & 0.000 & 0.588 & 0.262 & 0.336 \\ -0.568 & 0.916 & 0.000 & 0.000 & 0.483 & 0.511 \\ 0.248 & 0.754 & 0.446 & 1.371 & 0.000 & 0.595 \\ -0.405 & 0.189 & -0.182 & 0.693 & -0.470 & 0.000 \end{bmatrix}$$

Compared with the GARP bounds in $\mathbf{B}^G$, the use of the Afriat envelope functions has led to a tightening of two of the upper bounds to the base-weighted welfare index. The GARP upper bound to $\log Q_{51}^{LA}$ is $B_{15}^G = 0.981$, while the tightest upper bound is provided by $B_{15}^A = 0.954$. Similarly, the GARP upper bound to $\log Q_{54}^{LA}$ is $B_{45}^G = 0.511$, while the tightest upper bound is provided by $B_{45}^A = 0.483$.

The matrix of Afriat envelope upper bounds to the current-weighted welfare index is:

$$\mathbf{C}^A = \begin{bmatrix} 0.000 & 1.232 & 0.357 & 0.916 & 0.827 & 1.232 \\ -0.713 & 0.000 & -0.475 & 0.203 & -0.595 & -0.189 \\ -0.241 & 0.875 & 0.000 & 0.675 & 0.470 & 0.875 \\ -0.511 & 0.875 & 0.000 & 0.000 & 0.470 & 0.875 \\ 1.099 & 1.099 & 0.916 & 2.015 & 0.000 & 0.693 \\ -0.385 & 0.251 & -0.154 & 0.531 & -0.405 & 0.000 \end{bmatrix}$$

Compared with the GARP upper bounds to the current-weighted welfare index in $\mathbf{C}^G$, 15 of the upper bounds in $\mathbf{C}^A$ are tightened. Further, some of the bounds in $\mathbf{C}^A$ are significantly tighter than those in $\mathbf{C}^G$. For example, the GARP upper bound to $\log Q_{41}^{PA}$ is $C_{14}^G = 1.386$, while the tightest upper bound is provided by $B_{14}^A = 0.827$.

The above finding that the bilateral bounds constructed using the Afriat envelope approach are tighter than the GARP bounds is somewhat surprising since both approaches use RP relations and neither uses expansion path information (unlike the approach for constructing improved GARP bounds).

2.5 Conclusions

In this chapter, a thorough review of the revealed preference approach to welfare measurement has been presented. The first part of the chapter reviewed the testing of the representative consumer hypothesis. If the hypothesis of common preferences is not rejected for a given finite data set, then it is possible to construct bilateral true welfare indexes. Three methods for constructing bilateral true welfare indexes - the GARP bounds (Varian (1982)), improved GARP bounds (Blundell, Browning, and Crawford (1998)) and Afriat envelope bounds (Chavas and Cox (1997)) - were reviewed, and a framework for empirically comparing these methods was developed.

For an example data set, it was found that the improved GARP bounds were the tightest bounds to the bilateral true welfare indexes. However, this was an expected result, since this method incorporates expansion path information into the calculation of the true indexes. An unexpected finding was that of the other two methods reviewed (which do not use expansion path information), the Afriat envelope bounds were the tightest.

Chapter 3

Multilateral true indexes

3.1 Introduction

It was shown in Chapter 1 that homotheticity is necessary and sufficient for the existence of a unique (i.e. base-country invariant) welfare index. However, the unobservability of preferences means that the unique welfare index is indeterminate (but it is contained within the Paasche-Laspeyres (P-L) bounds). A unique true index was defined as a set of numbers, the ratios of the elements of which are contained with the P-L bounds. It was further shown that a unique true index, if it exists, will have the property of circularity, and thus it can be called a multilateral true index. This chapter focuses on the existence and construction of multilateral true welfare indexes.

In section 3.2, the RP test of common homothetic preferences, as developed by Afriat (1972, 1981), Diewert (1973) and Varian (1983) is reviewed. There are three ways the RP test of common homothetic preferences can be conducted, but the most easily implemented test is that of the Homothetic Axiom of Revealed Preference (HARP). It is shown that a test of HARP is equivalent to a test for the existence of a multilateral true index.

For a given data set that satisfies HARP there will in fact exist a infinite number of multilateral true indexes. However, one multilateral true index, the Ideal Afriat Index of Dowrick and Quiggin (1997), has certain desirable properties which make it appropriate for conducting cross-country comparisons of welfare. The Ideal Afriat Index is constructed for 1980 and 1993 ICP data and compared with another multilateral index (the EKS index).

The structure of this chapter is as follows. In Section 3.2, the RP tests of common homothetic preferences are reviewed. In Section 3.3, there is a review of the construction of multilateral true indexes, with particular focus on the Ideal Afriat Index. The Ideal Afriat Index is constructed for 1980 and 1993 ICP data in Sections 3.4 and 3.5, respectively. Conclusions are presented in Section 3.6.

3.2 RP Tests of Common Homothetic Preferences

Before stating the theory relating to the homothetic version of the RP test of common preferences, it is first discussed how homotheticity simplifies the Afriat inequalities (2.1) and (2.2). Using Roy's identity (see Appendix B), the Marshallian demand for good l is derived:

$$q_l(x, \mathbf{p}^i) = \frac{\frac{\partial \psi(\mathbf{p}^i)}{\partial p_l} x}{\psi(\mathbf{p}^i)} = q_l(1, \mathbf{p}^i)x,$$

or in matrix notation:

$$\mathbf{q}(x, \mathbf{p}^i) = \mathbf{q}(1, \mathbf{p}^i)x.$$

Therefore, we have:

$$\psi(x, \mathbf{p}^i) = u(\mathbf{q}(x, \mathbf{p}^i)) = u(\mathbf{q}(1, \mathbf{p}^i)x) = u(\mathbf{q}(1, \mathbf{p}^i))x, \tag{3.1}$$

since $u()$ is homogeneous of degree 1. By the Envelope Theorem (see Appendix B):

$$\frac{\partial \psi(x, \mathbf{p}^i)}{\partial x} = \Lambda^i,$$

where Λ^i is the (unobservable) marginal utility of income for country i. However, by (3.1), we also know that:

$$\frac{\partial \psi(x, \mathbf{p}^i)}{\partial x} = U^i,$$

and hence it must be the case that $U^i = \Lambda^i$. For the Afriat numbers to be consistent with homothetic utility maximisation, it must therefore be that $W^i = \lambda^i$. The Afriat inequalities (2.1) and (2.2) therefore become $W^i \leq W^j + W^j \mathbf{p}^j.(\mathbf{q}^i - \mathbf{q}^j)$, which simplifies to (denoting $A^i = W^i$ as the homothetic Afriat number):[1]

$$A^i \leq A^j \mathbf{p}^j.\mathbf{q}^i / \mathbf{p}^j.\mathbf{q}^j. \tag{3.2}$$

3.2.1 RP tests of common homothetic preferences - theory

The homothetic version of the RP test of the representative consumer hypothesis, as developed by Afriat (1972, 1981), Diewert (1973) and Varian (1983), tests the joint hypotheses that preferences are common and homothetic.

[1] As shown in Chapter 4, geometrically, homotheticity forces the Afriat outer-bound envelope functions through the origin.

Proposition 3.1. Common homothetic preferences - Afriat (1972, 1981), (Diewert 1973) and (Varian 1983) *The following conditions are equivalent:*[2]

i. *There exists a locally non-satiated homothetic utility function* $a(\mathbf{q})$ *that rationalises the data;*

ii. *The data satisfy the Homothetic Axiom of Revealed Preference (HARP): for all distinct choices of indexes* $(i, j, ..., m)$ *we have* $\left(\frac{\mathbf{p}^i.\mathbf{q}^j}{\mathbf{p}^i.\mathbf{q}^i}\right)\left(\frac{\mathbf{p}^j.\mathbf{q}^k}{\mathbf{p}^j.\mathbf{q}^j}\right)\ldots\left(\frac{\mathbf{p}^m.\mathbf{q}^i}{\mathbf{p}^m.\mathbf{q}^m}\right) \geq 1$*;*

iii. *There exists numbers* $A^i > 0$, $i = 1, \ldots, N$, *such that:*

$$A^i/A^j \leq \mathbf{p}^j.\mathbf{q}^i/\mathbf{p}^j.\mathbf{q}^j; \tag{3.3}$$

iv. *The numbers* $\exp(M_{ji}) = \min_{(j,k,\ldots,m,i)} \left\{\left(\frac{\mathbf{p}^j.\mathbf{q}^k}{\mathbf{p}^j.\mathbf{q}^j}\right)\left(\frac{\mathbf{p}^k.\mathbf{q}^l}{\mathbf{p}^k.\mathbf{q}^k}\right)\ldots\left(\frac{\mathbf{p}^m.\mathbf{q}^i}{\mathbf{p}^m.\mathbf{q}^m}\right)\right\}$, $i, j = 1, \ldots, N$, *exist;*

v. *There exists a concave, monotonic, continuous, locally non-satiated, homothetic utility function* $a(\mathbf{q})$ *that rationalises the data.*

Proof that (i)⇒(ii): For $(i, j, \ldots m)$ let:

$$\begin{aligned}
s^j &= \mathbf{v}^i.\mathbf{q}^j \\
s^k &= (\mathbf{v}^i.\mathbf{q}^j)(\mathbf{v}^j.\mathbf{q}^k) = s^j(\mathbf{v}^j.\mathbf{q}^k) \\
&\vdots \\
s^m &= (\mathbf{v}^i.\mathbf{q}^j)(\mathbf{v}^j.\mathbf{q}^k)\ldots(\mathbf{v}^l.\mathbf{q}^m) = s^l(\mathbf{v}^l.\mathbf{q}^m)
\end{aligned}$$

Note that use is made here of the normalised prices $\mathbf{v}^i = \mathbf{p}^i/(\mathbf{p}^i.\mathbf{q}^i)$, where $\mathbf{v}^i.\mathbf{q}^i = 1$ and $\mathbf{v}^i.(\mathbf{q}^j/s^j) = 1$. Since $a(\mathbf{q})$ rationalises the data, this implies that $a(\mathbf{q}^i) \geq a(\mathbf{q}^j/s^j)$ (because $\mathbf{q}^i$ was chosen when $\mathbf{q}^j/s^j$ was affordable).

Similarly, $(s^j\mathbf{v}^j).(\mathbf{q}^j/s^j) = 1$ and $(s^j\mathbf{v}^j).(\mathbf{q}^k/s^k) = 1$. Since by hypothesis $a(\mathbf{q})$ is homothetic and $\mathbf{q}^j$ is optimal at $\mathbf{v}^j$, this implies that $(\mathbf{q}^j/s^j)$ must be optimal at prices

[2] The presentation of this theorem and its proof is based on Varian (1983), however Varian (1983) did not include condition (iv). While condition 4 is not essential for showing that the test of HARP is equivalent to a test for common homothetic preferences (which was the main motivation behind Varian's presentation of the theorem), it is useful to include here since it introduces the bounds $\{M_{ij}\}$ to the true welfare index.

$(s^j\mathbf{v}^j)$ (the budget line has been shifted inward parallel and all optimal quantities are consequently divided by s^j). Therefore $a(\mathbf{q}^j/s^j) \geq a(\mathbf{q}^k/s^k)$ (since $\mathbf{q}^k/s^k$ is affordable at prices $s^j\mathbf{v}^j$.

By the same reasoning, $(s^l\mathbf{v}^l).(\mathbf{q}^l/s^l) = 1$ and $(s^l\mathbf{v}^l).(\mathbf{q}^m/s^m) = 1$. Hence $a(\mathbf{q}^l/s^l) \geq a(\mathbf{q}^m/s^m)$.

It has therefore been shown that $a(\mathbf{q}^i) \geq a(\mathbf{q}^j/s^j)$, $a(\mathbf{q}^j/s^j) \geq a(\mathbf{q}^k/s^k)$, and $a(\mathbf{q}^l/s^l) \geq a(\mathbf{q}^m/s^m)$. Combining these, $a(\mathbf{q}^i) \geq a(\mathbf{q}^m/s^m)$.

Suppose that $\mathbf{v}^m.(s^m\mathbf{q}^i) < \mathbf{v}^m.\mathbf{q}^m$. Since $a(\mathbf{q})$ exhibits non-satiation, this implies that $a(s^m\mathbf{q}^i) < a(\mathbf{q}^m)$. But by homotheticity, then it must be that $a(\mathbf{q}^i) < a(\mathbf{q}^m/s^m)$ (movement along rays from the origin). However, this contradicts $a(\mathbf{q}^l/s^l) \geq a(\mathbf{q}^m/s^m)$, which was shown above.

Therefore it must be that $\mathbf{v}^m.(s^m\mathbf{q}^i) \geq \mathbf{v}^m.\mathbf{q}^m = 1$ and therefore $s^m\mathbf{v}^m.\mathbf{q}^i \geq 1$, which is HARP.□

Proof that (ii)⇒(iii): Define

$$A^i = \min_{(j,k,\ldots,m,i)} \{(\mathbf{v}^j.\mathbf{q}^k)(\mathbf{v}^k.\mathbf{q}^l)\ldots(\mathbf{v}^m.\mathbf{q}^i)\}.$$

That is, A^i is a minimum of the given expression over all paths starting *anywhere* and terminating at i. It needs to be shown that this is well defined i.e. that a minimum actually exists. This is easily shown by use of an example. Consider the case where $N = 3$ and $i = 3$, then A^3 will be the minimum of the following expressions:

$$(\mathbf{v}^1.\mathbf{q}^2)(\mathbf{v}^2.\mathbf{q}^3)$$

$$(\mathbf{v}^1.\mathbf{q}^3)$$

$$(\mathbf{v}^2.\mathbf{q}^1)(\mathbf{v}^1.\mathbf{q}^3)$$

$$(\mathbf{v}^2.\mathbf{q}^3).$$

One may ask the question: why is A^i calculated as the minimum only over expressions which do not include *cycles* (a path which starts and finishes at the same point)? For example, why is $(\mathbf{v}^1.\mathbf{q}^2)(\mathbf{v}^2.\mathbf{q}^1)(\mathbf{v}^1.\mathbf{q}^3)$, which is one of the infinite number of paths which contain cycles, not included in the above paths over which A^i is calculated? The reason is that by HARP we only need to consider strings $(j, k, \ldots, m, i)$ which do not include cycles (since with HARP satisfied, any cycle must be ≥ 1) and hence its inclusion will either increase or have no effect on the expression. Using the above example, by HARP it must be that $(\mathbf{v}^1.\mathbf{q}^3) \leq (\mathbf{v}^1.\mathbf{q}^2)(\mathbf{v}^2.\mathbf{q}^1)(\mathbf{v}^1.\mathbf{q}^3)$ and hence the string with the cycle can be ignored. Given the data are finite, there are only a finite number of strings without cycles and therefore a minimum exists.

It is apparent that the A's defined in this way must be non-negative. What remains is to show that they satisfy the Afriat inequalities given in condition (iii). However it is readily apparent that: $A^i \leq A^j(\mathbf{v}^j.\mathbf{q}^i)$ since by definition A^i is the minimum over *all* paths to i.□

PROOF THAT (ii)⇒(iv). The number $\exp(M_{ji})$ is the minimum of the given expression over all paths starting at j and terminating at i. Given HARP is satisfied, it is possible to show that this minimum exists in the same way that it was shown that the numbers A^i exist.□

By definition, it must be that:

$$\exp(M_{ji}) \leq \mathbf{v}^j.\mathbf{q}^i, \tag{3.4}$$

since $\exp(M_{ji})$ is the minimum over all paths starting at j and terminating at i. However, the numbers $\exp(M_{ji})$ have a further special property that deserves highlighting. It is readily apparent that since: (a) A^i is the minimum of all (non-cyclical) paths starting anywhere and ending at i; (b) A^j is the minimum of all (non-cyclical) paths starting anywhere and ending at j and (c) $\exp(M_{ji})$ is the minimum of all (non-cyclical) paths starting at j and ending at i, then it must be the case that $A^i \leq A^j \exp(M_{ji})$. Hence, $A^i/A^j \leq \exp(M_{ji})$ and combining this with (3.4) it must be that:

$$A^i/A^j \leq \exp(M_{ji}) \leq \mathbf{v}^j.\mathbf{q}^i.$$

This result shows that $\exp(M_{ji})$ is the tightest upper bound to the unique true index A^i/A^j, and $\exp(M_{ji})$ is itself bounded by the Laspeyres quantity index. This result is used later in the construction of bounds to true welfare indexes.

PROOF THAT (iii)⇒(v): Define $a(\mathbf{q})$ as:

$$a(\mathbf{q}) = \min_i A^i \mathbf{v}^i.\mathbf{q}.$$

It is clear that this function has the stated properties. It can be verified that it rationalises the data using an approach similar to the proof similar to that in Proposition 2.1. First note that $a(\mathbf{q}^i) = A^i$ for all $i = 1, \ldots, N$. Suppose instead that the minimum is obtained at $\mathbf{q}^m$, then:

$$a(\mathbf{q}^i) = A^m \mathbf{v}^m.\mathbf{q}^i \leq A^i,$$

but this contradicts the Afriat inequalities.

Suppose we have some $\mathbf{q}$ such that $\mathbf{p}^j.\mathbf{q}^j \geq \mathbf{p}^j.\mathbf{q}$. From the following set of inequalities, we can show that $a(\mathbf{q}^j) \geq a(\mathbf{q})$, as required:

$$\begin{aligned} a(\mathbf{q}) &= \min_i A^i \mathbf{v}^i.\mathbf{q} \\ &\leq A^j \mathbf{v}^j.\mathbf{q} \\ &\leq A^j = a(\mathbf{q}^j).\square \end{aligned}$$

PROOF THAT (v)⇒(i): This is obvious.

As was the case for testing for common non-homothetic preferences, there are three approaches for testing for common homothetic preferences. The test of HARP is the easiest to implement, but it is useful to outline the other two approaches.

3.2.2 Three tests of common homothetic preferences

Testing HARP

By Proposition 3.1, a test that the data satisfy HARP is equivalent to a test of common homothetic preferences. It is convenient to restate HARP in the form: $\ln\left(\frac{\mathbf{p}^i.\mathbf{q}^j}{\mathbf{p}^i.\mathbf{q}^i}\right) + \ln\left(\frac{\mathbf{p}^j.\mathbf{q}^k}{\mathbf{p}^j.\mathbf{q}^j}\right) + \ldots + \ln\left(\frac{\mathbf{p}^m.\mathbf{q}^i}{\mathbf{p}^m.\mathbf{q}^m}\right) \geq 0$. Or, equivalently: $L_{ij} + L_{jk} + \ldots + L_{mi} \geq 0$. As with the test of GARP, testing HARP involves the use of Warshall's algorithm. Re-define $\mathbf{L}$ as an N by N matrix of "costs" of moving from location to location, with a particular element L_{ij} interpreted as the cost of moving from location i to location j. In this context, the HARP inequality is a test of whether the cost of moving from i to itself can be made cheaper than zero; testing for HARP is therefore a test of whether there are any negative cost cycles in the data. The test of HARP can be summarised as follows: use Warshall's algorithm to replace each L_{ij} with the minimum cost of moving from node i to j, thus constructing the minimum path matrix $\mathbf{M}$:[3]

$$M_{ij} = \min_{k,\dots,m} \{L_{ij}, (L_{ik} + L_{kl} + \ldots + L_{mj})\}.$$

If any of the diagonal elements of $\mathbf{M}$ are negative then this indicates a negative cost cycle and HARP is therefore violated. There is a more intuitive explanation for why the existence of homothetic preferences requires that HARP be satisfied. If HARP was not satisfied then it would be the case that $L_{ij} + L_{jk} + \ldots + L_{mi} < 0$. Since with homothetic preferences it must be that $a^j - a^i \leq L_{ij}$ (this is condition (iii) in Proposition 3.1), where $a^i = \log(A^i)$, then it must be equally true that,$(a^j - a^i) + (a^k - a^j) + \ldots + (a^i - a^m) < 0$. However, this implies that $(a^i - a^i) < 0$ and hence HARP must be satisfied for preferences to be homothetic.

Constructing the Afriat numbers using combinatorial methods

The second method for checking whether a particular set of data is consistent with common homothetic preferences involves calculating the Afriat numbers directly using a modified version of Algorithm 2.1.

Algorithm 3.1. CONSTRUCTING HOMOTHETIC AFRIAT NUMBERS *Input: A set of demand observations* $(\mathbf{p}, \mathbf{q})$ *and the revealed preference relation R that satisfies HARP. An algorithm* $\max(I)$ *which finds the maximal element from a set of demand observations I. Output: A set of numbers* $A^i > 0$, $i = 1, \ldots, N$ *that satisfy the Afriat inequalities* (3.3).

1. $I = 1, \ldots, N$, $B = \varnothing$.

2. *Let* $m = \max(I)$.

3. *Set* $E = \{i \text{ in } I : \mathbf{q}^i R \mathbf{q}^m\}$. *If* $B = \varnothing$, *set* $A^i = 1$ *and go to 5. Otherwise go to 4.*

4. *Set* $A^m = \min_{i \in E} \min_{j \in B} \min\{A^j \mathbf{p}^j.\mathbf{q}^i / \mathbf{p}^j.\mathbf{q}^j), A^j\}$.

5. *Set* $A^i = A^m$ *for all* $i \in E$.

6. *Set* $I = I \setminus E$, $B = B \cup E$. *If* $I = \varnothing$, *stop. Otherwise go to 2.*

[3] It is important to note that in constructing the minimum path matrix all of the relevant $\mathbf{L}$'s could be added together or just a subset.

Constructing the Afriat numbers using mathematical programming

The mathematical programs of Diewert and Parkan (1985) which were used to test for common general preferences can be modified for testing for common homothetic preferences. Under the null hypothesis of homotheticity $W^i = \lambda^i = A^i$, and therefore (2.5) can be converted to the following unconstrained problem:

$$\min_{A,\sigma} \sum_{i \in N} \sum_{j \in N} [s^{ij}]^2 \quad \text{subject to}$$

$$s^{ij} \equiv A^j - A^i \mathbf{p}^i.\mathbf{q}^j / \mathbf{p}^i.\mathbf{q}^i + \sigma_{ij}^2 \quad i,j \in N$$

where, as before, $S^{ij} \equiv \sigma_{ij}^2 \geq 0$. If the objective function in the above problem attains its lower bound of zero, then there exists a homothetic utility function which rationalises the data. With normalised prices, this problem becomes:

$$\min_{A,\sigma} \sum_{i \in N} \sum_{j \in N} [s^{ij}]^2 \quad \text{subject to}$$

$$s^{ij} \equiv A^j - A^i \mathbf{v}^i.\mathbf{q}^j + \sigma_{ij}^2 \quad i,j \in N \tag{3.5}$$

3.2.3 Example

The three tests of common homothetic preferences were applied to the 6 country data introduced in Chapter 2 (Section 2.2.4).

Test of HARP

The minimum path matrix for the 6 country data is:

$$\mathrm{M} = \begin{bmatrix} 0.000 & 1.037 & 0.511 & 0.693 & 0.377 & 0.803 \\ -0.871 & -0.044 & -0.514 & -0.178 & -0.704 & -0.278 \\ -0.357 & 0.526 & 0.000 & 0.336 & -0.134 & 0.292 \\ -0.568 & 0.469 & -0.057 & 0.000 & -0.191 & 0.235 \\ -0.117 & 0.709 & 0.239 & 0.576 & 0.000 & 0.476 \\ -0.682 & 0.145 & -0.325 & 0.011 & -0.514 & -0.089 \end{bmatrix}$$

Since two of the elements on the diagonal are negative, these data do not satisfy HARP and hence cannot be rationalised by a homothetic utility function. The fact that these data satisfied GARP illustrates that HARP is in general more difficult to satisfy than

GARP, since it implies that the expansion paths be rays from the origin (while GARP places no restriction on the expansion paths). With country 6 dropped, the minimum path matrix is:

$$\mathbf{M} = \begin{bmatrix} 0.000 & 1.153 & 0.511 & 0.693 & 0.614 \\ -0.827 & 0.000 & -0.470 & -0.134 & -0.539 \\ -0.357 & 0.642 & 0.000 & 0.336 & 0.103 \\ -0.568 & 0.585 & -0.057 & 0.000 & 0.046 \\ -0.073 & 0.754 & 0.284 & 0.620 & 0.000 \end{bmatrix}$$

Since none of the elements of the diagonal of $\mathbf{M}$ are negative, the data satisfy HARP and hence can be rationalised by a homothetic utility function.

Constructing the Afriat numbers using combinatorial methods

Algorithm 3.1 was run using the 5 country data and the following Afriat numbers were calculated (thus indicating that the data can be rationalised by a homothetic utility function): $A^1 = 0.438$, $A^2 = 1$, $A^3 = 0.625$, $A^4 = 1$ and $A^5 = 0.438$. Note, however that these numbers are not unique (there will, in general, be an infinite combination of numbers which satisfy the Afriat inequalities), and hence they cannot be used for computing utility bounds.

Constructing the Afriat numbers using mathematical programming

The third RP test of whether the data can be rationalised by a homothetic utility function is conducted by solving the mathematical program (3.5). For the particular starting values used, and normalising $A^3 = 1$, the Afriat numbers calculated are: $A^1 = 0.6$, $A^2 = 1.644$, $A^3 = 1$, $A^4 = 1.114$ and $A^5 = 0.774$. As above, these Afriat numbers are not unique and hence do not provide any information on the bounds to the true welfare index.

3.3 Constructing Multilateral true indexes

A multilateral true index was defined in Definition 1.5 as a true index which also has the property of being a multilateral index (i.e. it satisfies circularity). The Afriat numbers constructed using mathematical programming and combinatorial methods are by definition multilateral true indexes. However, these numbers are dependent on the particular starting values used, and they do not contain any information on the bounds to the unique welfare index. In contrast, the Afriat numbers constructed during the HARP test of common homothetic preferences (and are elements of the minimum path matrix) provide information in that they are tight *bounds* to all possible true welfare indexes.

Dowrick and Quiggin (1997) have developed several general results on the construction of multilateral true indexes using the elements of the minimum path matrix, which are now summarised. In proving Proposition 3.1, it was shown that given that common homothetic preferences cannot be rejected by the data, then the system:

$$A^i/A^j \leq \exp(M_{ji}) \leq \mathbf{p}^j.\mathbf{q}^i/\mathbf{p}^j.\mathbf{p}^j \quad i,j \in N,$$

has a positive solution. By symmetry, it is also true (since $A^j/A^i \leq \exp(M_{ij})$ and hence $A^i/A^j \geq 1/\exp(M_{ij})$) that:

$$\mathbf{p}^i.\mathbf{q}^i/\mathbf{p}^i.\mathbf{p}^j \leq 1/\exp(M_{ij}) \leq A^i/A^j \leq \exp(M_{ji}) \leq \mathbf{p}^j.\mathbf{q}^i/\mathbf{p}^j.\mathbf{p}^j.$$

Using the logarithmic Afriat numbers a^i, the matrix of log upper-bounds $\mathbf{M}$ to the set $\mathcal{A}$ of bilateral true indexes is represented:

$$\mathcal{A} \equiv \{\mathbf{a} \equiv (a^1, a^2, \ldots, a^N) | (a^j - a^i) \leq M_{ij} \leq \ln(\mathbf{p}^i.\mathbf{q}^j/\mathbf{p}^i.\mathbf{q}^i) \equiv L_{ij}\}. \tag{3.6}$$

With homotheticity, $\mathbf{M}$ exists and it is in fact the minimum path matrix.[4] Using symmetry we can rewrite (3.6) as:

$$-M_{ji} \leq (a^j - a^i) \leq M_{ij} \quad i,j = 1, ..., N. \tag{3.7}$$

Any true index, $\mathbf{a}$, must satisfy these inequalities for each i, j. From the definition of the minimum path, it must be that:

$$M_{kj} \leq M_{ki} + M_{ij} \quad i,j = 1, ..., N. \tag{3.8}$$

[4]It follows from the above that with homotheticity $L_{ij} \geq a^j - a^i$ and $L_{ji} \geq a^i - a^j$ and hence $L_{ij} + L_{ji} \geq 0$; the Laspeyres index must be greater than the Paasche index.

From this, $M_{kj} - M_{ki} \leq M_{ij}$; the terms on the left of this inequality are the jth and ith elements of the kth row of $\mathbf{M}$; this row therefore constitutes a true index as defined by (3.7). The minimum path inequality (3.8) can also be written as $-M_{ij} + M_{kj} \leq M_{ki}$ and hence the negative elements of column j also comprise a true index.

These results can be extended to the following general proposition on the construction of multilateral true indexes.

Proposition 3.2. MULTILATERAL TRUE INDEXES; DOWRICK AND QUIGGIN (1997) *Let $f : \Re^N \to \Re$ be any function which is non-decreasing in each of its arguments and for which $f(\mathbf{x} + \mathbf{c}) = f(\mathbf{x}) + c$ where $\mathbf{c} = (c, c, ..., c)$, and let $\mathbf{M}_{.j}$ and $\mathbf{M}_{j.}$ denote the jth column and jth row respectively of the minimum path matrix, $\mathbf{M}$. $\{z^j | [z^j = f(\mathbf{M}_{.j}), j = 1, ..., N]$ or $[z^j = f(-\mathbf{M}_{j.}), j = 1, ..., N]\}$ is a true index.*

PROOF: It is apparent from (3.8) that:

$$\begin{aligned} M_{1j} &\leq M_{1i} + M_{ij} \\ M_{2j} &\leq M_{2i} + M_{ij} \\ &\vdots \\ M_{Nj} &\leq M_{ni} + M_{ij}. \end{aligned}$$

Hence, $f(\mathbf{M}_{.j}) \leq f(\mathbf{M}_{.i} + M_{ij}\mathbf{i}_N)$, where $\mathbf{i}_N$ is a $N \times 1$ vector of ones. From the definition of $f()$, $f(\mathbf{M}_{.j}) \leq f(\mathbf{M}_{.i}) + M_{ij}$, and therefore, $f(\mathbf{M}_{.j}) - f(\mathbf{M}_{.i}) \leq M_{ij}$ Similarly, from (3.8):

$$\begin{aligned} -M_{i1} - (-M_{k1}) &\leq M_{ki} \\ -M_{i2} - (-M_{k2}) &\leq M_{ki} \\ &\vdots \\ -M_{in} - (-M_{kn}) &\leq M_{ki}. \end{aligned}$$

Therefore, $f(-\mathbf{M}_{i.} - M_{ki}\mathbf{i}_N) \leq f(-\mathbf{M}_{k.})$, $f(-\mathbf{M}_{i.}) - M_{ki} \leq f(-\mathbf{M}_{k.})$ and hence $f(-\mathbf{M}_{i.}) - f(-\mathbf{M}_{k.}) \leq M_{ki}$. We therefore have that $z^j - z^i \leq M_{ij}$ for all i, j and from (3.6), $\mathbf{z} \in \mathcal{A}$. □

The elements of a true index can therefore be constructed by applying the appropriate transformation f to each column or to each (negative) row of the minimum path matrix.

Examples of f are the ith element, or any other rank-dependent functions such as the minimum, maximum and median.[5]

From the definition of the true index in (3.6), any convex combination of true indexes is also a true index. To see this, assume we have two sets of bilateral true indexes $\mathbf{a} \equiv \{a^1, a^2, \ldots, a^N | (a^j - a^i) \leq M_{ij}\}$ and $\mathbf{b} \equiv \{b^1, b^2, \ldots, b^N | (b^j - b^i) \leq M_{ij}\}$. Define the convex combination of $\mathbf{a}$ and $\mathbf{b}$ as $\mathbf{c} \equiv [\kappa a^1 + (1-\kappa)b^1, a^2 + (1-\kappa)b^2, \ldots, a^N + (1-\kappa)b^N]$ where $\kappa \in [0, 1]$. To see that $\mathbf{c}$ is a true index, consider the following inequalities:

$$\begin{aligned} \kappa[a^j - a^i] &\leq \kappa M_{ij} \\ (1-\kappa)[b^j - b^i] &\leq (1-\kappa) M_{ij} \end{aligned}$$

The sum of these inequalities is $[\kappa a^j + (1-\kappa)b^j] - [\kappa a^i + (1-\kappa)b^i] \leq M_{ij}$ and we can therefore conclude that $\mathbf{c}$ is indeed a true index.

The following are therefore multilateral true indexes:

1. any row of the minimum path matrix, $\mathbf{M}$, or of its negative transpose $-\mathbf{M}^T$;

2. any row of the transformed matrix, $\mathbf{M}^r$, or $-\mathbf{M}^{Tr}$, where each column is the corresponding column of $\mathbf{M}$ or $-\mathbf{M}^T$ reordered by rank;

3. any convex combination of the above indexes.

3.3.1 Graphical example of the minimum path algorithm

From Result 1.3 we know that with homotheticity there exists a unique welfare index which is bounded by the Laspeyres and Paasche indexes. It is now graphically shown how multilateral comparisons (and the minimum path algorithm in particular) can be used to tighten the bounds to the true welfare index. Assume we are wanting to compare the welfare between two countries for which the Laspeyres index exceeds the Paasche index (see Figure 3.2).

[5] A potential source of confusion here is that previously it was stated that the elements of a row and (negative) column of $\mathbf{M}$ are true indexes, but in Proposition 3.2 it is stated that $f(\mathbf{M}_{.j})$ (a function of column j of $\mathbf{M}$) and $f(-\mathbf{M}_{j.})$ (a function of the negative of row j of $\mathbf{M}$) comprise true indexes. The explanation is that $f(\mathbf{M}_{.j})$ is reducing each column to a single element and the result of this is a *row* of numbers which is a true index. Similarly, $f(-\mathbf{M}_{j.})$ is reducing each row to one element and this results in a *column* of numbers which is a true index.

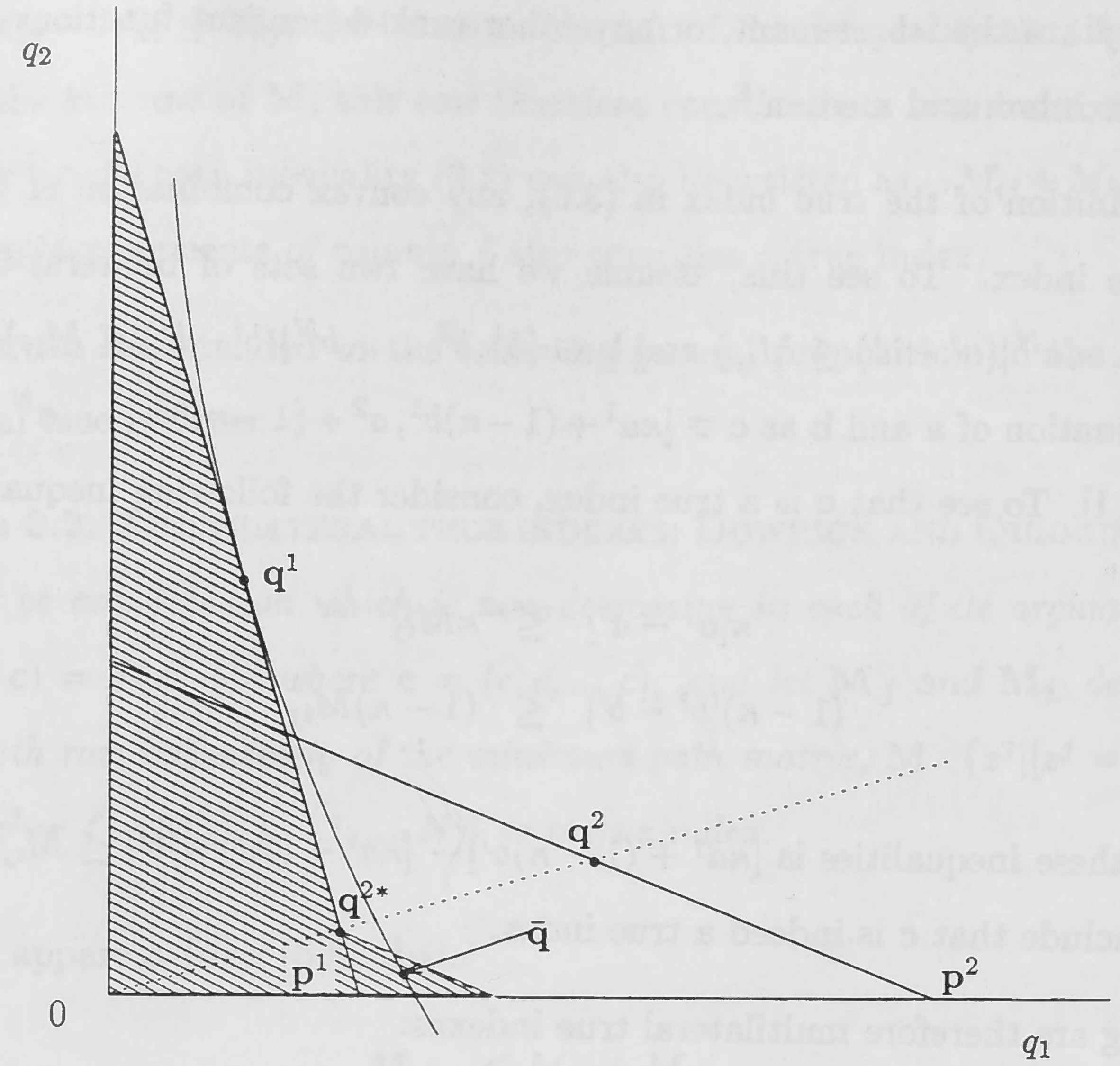

Figure 3.1: Constructing a lower bound to country 1's indifference curve - $N = 2$

Since these data satisfy homotheticity (as shown in Proposition 1.3, with $N = 2$ the P-L in equality is necessary and sufficient for this) we know the expansion paths are rays from the origin and we can use these paths to construct bounds to the indifference curves. An argument similar to that used in Section 2.3 can be employed to show that country 1's indifference curve cannot be in the shaded region in Figure 3.1. In particular, if the indifference curve *were* in this region, then we could have the following: $\mathbf{q}^{2*} R^D \bar{\mathbf{q}}$ and $\bar{\mathbf{q}} R^D \mathbf{q}^{2*}$, that is, a violation of WARP. Hence it must be that with homotheticity, the shaded region forms a lower bound to country 1's indifference curve. This lower bound is in fact the *unconditional homothetic outer bound* to country 1's indifference curve.[6]

The lower bound to country 2's indifference curve can similarly be constructed (Figure 3.2). These lower bounds can be used to construct bounds to the welfare indexes which are tighter than those provided by the fixed-weight and classical bounds. In particular, it is apparent that $\mathbf{p}^1.\mathbf{q}^{1*}$ is an improved lower bound to $e(U^2, \mathbf{p}^1)$ and $\mathbf{p}^2.\mathbf{q}^{2*}$ is an

[6]See Chapter 4 for more details on the construction of this bound.

improved lower bound to $e(U^1, \mathbf{p}^2)$ and hence:

$$\frac{\mathbf{p}^1.\mathbf{q}^{1*}}{\mathbf{p}^1.\mathbf{q}^1} \leq \frac{e(U^2, \mathbf{p}^1)}{e(U^1, \mathbf{p}^1)} \leq \frac{\mathbf{p}^1.\mathbf{q}^2}{\mathbf{p}^1.\mathbf{q}^1}$$
$$\frac{\mathbf{p}^2.\mathbf{q}^2}{\mathbf{p}^2.\mathbf{q}^1} \leq \frac{e(U^2, \mathbf{p}^2)}{e(U^1, \mathbf{p}^2)} \leq \frac{\mathbf{p}^2.\mathbf{q}^2}{\mathbf{p}^2.\mathbf{q}^{2*}}$$

However, it is immediately apparent from the diagram that:

$$\frac{\mathbf{p}^1.\mathbf{q}^{1*}}{\mathbf{p}^1.\mathbf{q}^1} = \frac{0a}{0b} = \frac{0c}{0d} = \frac{\mathbf{p}^2.\mathbf{q}^2}{\mathbf{p}^2.\mathbf{q}^1},$$

that is, with $N = 2$ the lower bound to country 2's indifference curve is redundant in the construction of a lower bound to Q_{21}^{LA}. Similarly, it can be shown that with $N = 2$ the lower bound to country 1's indifference curve is redundant in the construction of an upper bound to Q_{21}^{PA}. Thus, we have graphically shown the result in Chapter 1 that with $N = 2$ the tight bound to the welfare index is provided by the Paasche and Laspeyres indexes.

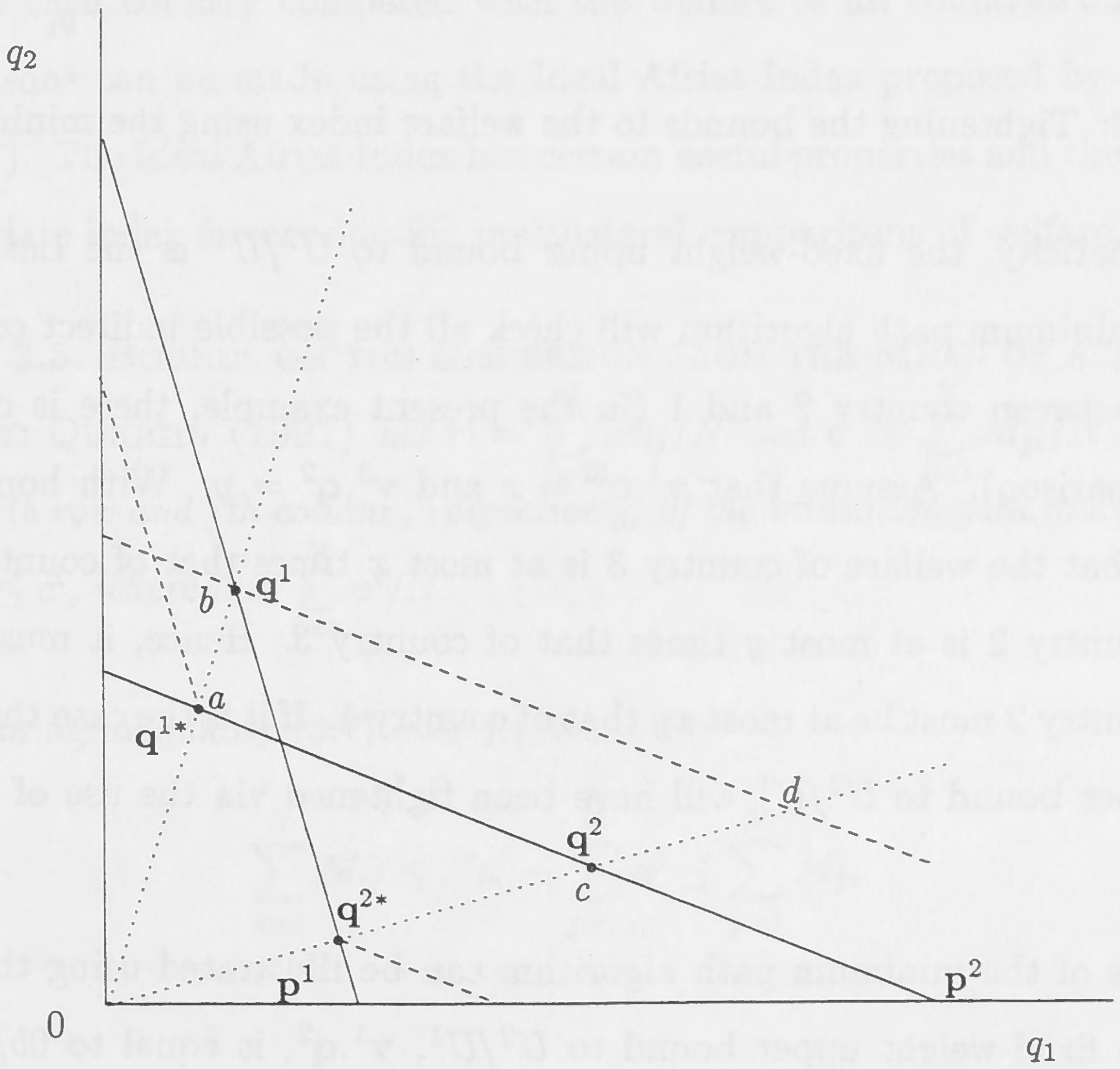

Figure 3.2: Constructing bounds to the true index - $N = 2$

The lower bounds to the indifference curves, and (relatedly), the bounds to the welfare index comparing countries 1 and 2 can be improved by the introduction of a third country that shares common homothetic preferences with countries 1 and 2 (see Figure 3.3).

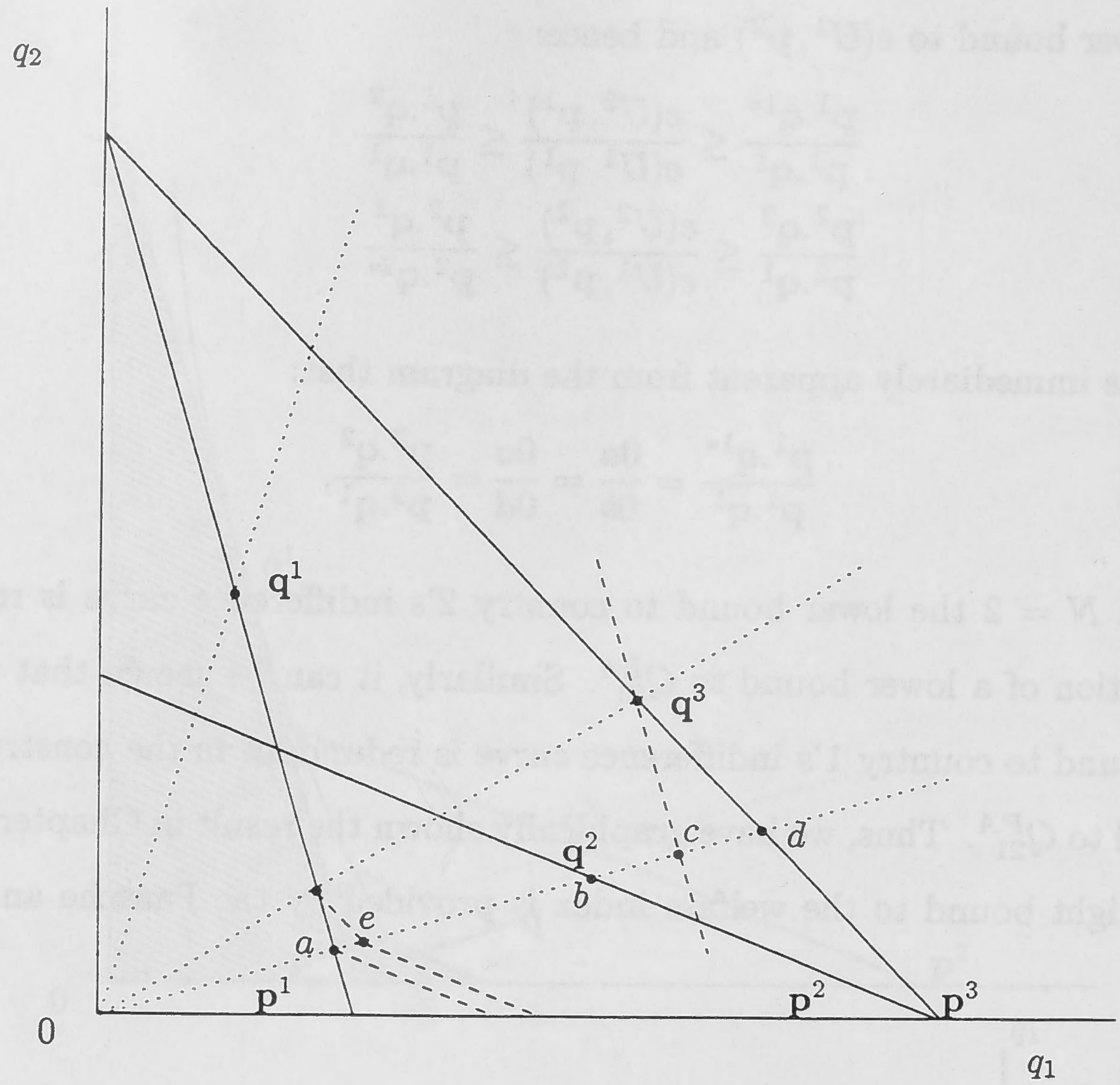

Figure 3.3: Tightening the bounds to the welfare index using the minimum path

With homotheticity, the fixed-weight upper bound to U^2/U^1 is the Laspeyres index, $\mathbf{v}^1.\mathbf{q}^2$. The minimum path algorithm will check all the possible indirect comparisons of the welfare between country 2 and 1 (in the present example, there is only one such indirect comparison). Assume that $\mathbf{v}^1.\mathbf{q}^3 = x$ and $\mathbf{v}^3.\mathbf{q}^2 = y$. With homotheticity, it is apparent that the welfare of country 3 is at most x times that of country 1, and the welfare of country 2 is at most y times that of country 3. Hence, it must be that the welfare of country 2 must be at most xy that of country 1. If it is the case that $xy < \mathbf{v}^1.\mathbf{q}^2$ then the upper bound to U^2/U^1 will have been tightened via the use of third-country comparisons.

The workings of the minimum path algorithm can be illustrated using the diagram as follows. The fixed-weight upper bound to U^2/U^1, $\mathbf{v}^1.\mathbf{q}^2$, is equal to $0b/0a$. Further, $\mathbf{v}^1.\mathbf{q}^3 = 0c/0a = 0d/0e$ (where the latter equality is found using similar triangles) and $\mathbf{v}^3.\mathbf{q}^2 = 0b/0d$. The minimum path is $(\mathbf{v}^1.\mathbf{q}^3)\ (\mathbf{v}^3.\mathbf{q}^2) = 0d/0e.0b/0d = 0b/0e < 0b/0a$ and thus the upper bound to U^2/U^1 has been tightened via the inclusion of country 3 into the analysis.

3.3.2 Bilateral comparisons - example

As discussed above, if the data are consistent with homotheticity, then the tightest bounds to the unique welfare index are provided by the elements of **M**. The elements of **M** can therefore be used to make multilaterally consistent bilateral comparisons between observations. From the minimum path matrix constructed for the 5 country data (which satisfy HARP) in Section 3.2.3, we can therefore say that the welfare of country 5 is at most $e^{0.103} = 1.11$ and at least $e^{-0.284} = 0.75$ times the welfare of country 3.

3.3.3 Multilateral comparisons - the Ideal Afriat Index

Above it was shown how the elements of **M** can be used to make multilaterally consistent bilateral comparisons between observations.[7] However, in a cross-country context, one generally wants to make multilaterally consistent multilateral comparisons of welfare; i.e. the welfare of each country compared with the welfare of all countries in the sample. Such comparisons can be made using the Ideal Afriat Index proposed by Dowrick and Quiggin (1997). The Ideal Afriat Index has certain useful properties and thus is proposed as an appropriate index for conducting multilateral comparisons of welfare.

Proposition 3.3. Bounds on the Dispersion from the Mean of a True Index; Dowrick and Quiggin (1997) *Let* $r^i = \sum_{j=1}^{N} M_{ij}/N$ *and* $c^i = \sum_{j=1}^{N} M_{ji}/N$ *represent the means of the* i*th row and* j*th column, respectively, of the minimum path matrix,* $\mathbf{M}$. *Then* $-r^i \leq a^i - \bar{a} \leq c^i$, *where* $\bar{a} \equiv \sum_{i=1}^{N} a^i/N$.

Proof: Summing inequality (3.7) over j gives:

$$\sum_{j=1}^{N} M_{ij} \leq Na^i - \sum_{j=1}^{N} a^j \leq \sum_{j=1}^{N} M_{ji}$$

or,

$$-r^i \leq a^i - \bar{a} \leq c^i. \quad \square \tag{3.9}$$

[7] A cross-country example of multilaterally consistent bilateral comparisons would be using **M** to compare the welfare of all countries in the sample to that of the U.S., for example. Another (more common) example of such comparisons is in the context of constructing price indexes where **M** is used to compare the cost of living in each year to that of a particular base year; for examples see Varian (1983) and Manser and McDonald (1988).

It can be shown that the bounds in (3.9) are tight in that they are attainable using a combination of elements from $\mathbf{M}$. The true index $\mathbf{a} = \mathbf{M}_{i.}$ gives the lower bound since its mean value is, by definition r^i, $a^i = M_{ii} = 0$ and therefore $a^i - \bar{a} = 0 - r^i = -r^i$. Similarly, the true index $\mathbf{a} = -\mathbf{M}_{.i}$ gives the upper bound since its mean value is $-c^i$, $a^i = -M_{ii} = 0$ and therefore $a^i - \bar{a} = 0 - (-c^i) = c^i$.

Proposition 3.4. The Ideal Afriat Index; Dowrick and Quiggin (1997) *Let* $\mathbf{a}^+ \equiv (c^1, c^2, \ldots, c^N)$ *and* $\mathbf{a}^- \equiv (-r^1, -r^2, \ldots, -r^N)$ *represent the vector of column means and the vector of negative row means, respectively, of the minimum path matrix,* $\mathbf{M}$. *Define the Ideal Afriat Index as* $\mathbf{a}^* \equiv (\mathbf{a}^+ + \mathbf{a}^-)/2$, *the vector of overall means.*

(i) $\mathbf{a}^*$, $\mathbf{a}^+$ *and* $\mathbf{a}^-$ *are true indexes;*

(ii) the numbers a^{+i} *and* a^{-i} *are the upper and lower bounds on the true index number for observation i relative to the sample mean;*

(iii) a^{*i} *is the midpoint of the range of* $a^i - \bar{a}$;

(iv) $\mathbf{a}^*$ *is equivalent to the Ideal Fisher Index when* $N = 2$.

Proof: From Proposition 3.2, any row of $\mathbf{M}$ is a true index, as is the negative of any column vector and its convex combination. Statement (i) follows from this and (ii) and (iii) follow from (3.9) and the construction of a^{*i}. When there are only two observations $\mathbf{M}$ is equal to $\mathbf{L}$ and it follows that $\mathbf{a}^*$ is equivalent to the Ideal Fisher Index.

3.3.4 The Ideal Afriat Index: example

The Ideal Afriat Index is now constructed for the 5 country data which satisfy homotheticity. For comparative purposes, the Ideal Afriat Index is compared with another multilateral index, the EKS index (as discussed in Section 1.5, however, the EKS index is not a multilateral true index).

From Proposition 3.2, the ith row of $\mathbf{M}$, $M_{i.}$, is a true index, as is the negative of the ith column of $\mathbf{M}$, $-M_{.i}$. Further, the midpoint of these two indexes which is the vector $(M_{i.} - M_{.i})/2$ will also be a true index. Applying this to $\mathbf{M}$ calculated for the 5 country data (Section 3.2.3) gives the following matrix of midpoint true indexes (note that this is equivalent to constructing Fisher indexes from $\mathbf{M}$ rather than from $\mathbf{L}$):

$$\overline{\mathbf{M}} = \begin{bmatrix} 0.000 & 0.990 & 0.434 & 0.631 & 0.343 \\ -0.990 & 0.000 & -0.556 & -0.359 & -0.646 \\ -0.434 & 0.556 & 0.000 & 0.197 & -0.090 \\ -0.631 & 0.359 & -0.197 & 0.000 & -0.287 \\ -0.343 & 0.646 & 0.090 & 0.287 & 0.000 \end{bmatrix}$$

Thus, the row $\overline{M}_{3.}$ gives a true index comparing the welfare of all countries with that of country 3 which is the midpoint of bounds to the true index. For example, it was shown above that the welfare of country 5 is at most 1.11 and at least 0.75 times the welfare of country 3, the midpoint indicates the welfare of country 5 is $e^{-0.090} = 0.91$ times the welfare of country 3.[8]

It was shown in Chapter 1 that a true index satisfies the property of circularity, that is, a welfare comparison between countries i and j must be consistent with a welfare comparison constructed via third country k. The property of circularity is evident in the true indexes in $\overline{M}$. For example, $\overline{M}_{35} = -0.09$ and $\overline{M}_{31} + \overline{M}_{15} = -0.434 + 0.343 = -0.09$. The matrix $\overline{M}$ can be used to make multilaterally consistent bilateral comparisons (i.e. between country i and j), but in cross country analysis one is generally more interested in making multilaterally consistent multilateral comparisons (i.e. comparing the welfare of country i with that of the sample mean). The Ideal Afriat Index can be calculated as the mean of the columns of $\overline{M}$, or equivalently, the negative of the mean of the rows of $\overline{M}$. From Table 3.1, the welfare of country 5 is between 27.2 percent ($1 - e^{-0.317} = 0.728$) below and 4.6 percent above the sample mean, with a midpoint estimated welfare of 87.3 percent of the sample mean.

Table 3.1: Ideal Afriat Index and bounds, 5 country data

	Country 1	Country 2	Country 3	Country 4	Country 5
$\mathbf{a}^*$	−0.479	0.510	−0.046	0.151	−0.136
$\mathbf{a}^-$	−0.594	0.394	−0.145	−0.001	−0.317
$\mathbf{a}^+$	−0.365	0.627	0.053	0.303	0.045

[8] i.e. $\overline{M}_{35} = (M_{35} - M_{53})/2$.

Comparison with a multilateral index: the EKS index

The matrix of log Laspeyres indexes for countries 1-5 is:

$$\mathbf{L} = \begin{bmatrix} 0.000 & 1.365 & 0.511 & 0.693 & 1.126 \\ -0.736 & 0.000 & -0.470 & 0.318 & -0.539 \\ -0.357 & 0.642 & 0.000 & 0.588 & 0.262 \\ -0.568 & 0.916 & 0.000 & 0.000 & 0.710 \\ 0.248 & 0.754 & 0.446 & 1.371 & 0.000 \end{bmatrix}$$

The basic "building block" of the EKS index is the Fisher index: $Q_{ij}^F = \sqrt{Q_{ij}^L Q_{ij}^P}$. For the present example, the matrix of logarithms of Fisher indexes is:

$$\mathbf{F} = \begin{bmatrix} 0.000 & 1.050 & 0.434 & 0.631 & 0.439 \\ -1.050 & 0.000 & -0.556 & -0.299 & -0.646 \\ -0.434 & 0.556 & 0.000 & 0.294 & -0.092 \\ -0.631 & 0.299 & -0.294 & 0.000 & -0.330 \\ -0.439 & 0.646 & 0.092 & 0.330 & 0.000 \end{bmatrix}$$

While the Fisher index has many desirable properties, it does not have the property of circularity. The lack of circularity of the Fisher index is evident from the elements of $\mathbf{F}$, for example, $F_{35} = -0.09$ and $F_{31} + F_{15} = -0.434 + 0.439 = 0.01$. The non-circularity of the Fisher index makes it unsuitable for use in multilateral comparisons, however the EKS index, which is a multilateral generalisation of the Fisher index, is commonly used in cross country welfare comparisons. The EKS index is the geometric mean of the ratios of all N bilateral Fisher indexes, taking each country in turn as base: $Q_{ij}^{EKS} = \prod_{k=1}^{N} \left(\frac{Q_{ik}^F}{Q_{jk}^F} \right)^{(1/N)}$. The matrix of logarithms of EKS indexes for the present example is:

$$\mathbf{E} = \begin{bmatrix} 0.000 & 1.021 & 0.446 & 0.702 & 0.385 \\ -1.021 & 0.000 & -0.575 & -0.319 & -0.636 \\ -0.446 & 0.575 & 0.000 & 0.256 & -0.061 \\ -0.702 & 0.319 & -0.256 & 0.000 & -0.317 \\ -0.385 & 0.636 & 0.061 & 0.317 & 0.000 \end{bmatrix}$$

The EKS index satisfies the property of circularity: $E_{35} = -0.06$ and $E_{31} + E_{15} = -0.446 + 0.385 = 0.06$. However, as noted in Section 1.5.2 the circularity of the EKS index is purely a statistical artifact; it arises from the method of calculating the EKS (averaging over all countries), not because the underlying preference structure necessarily imparts circularity. While the EKS index satisfies circularity by construction, it is not a true index as defined in Definition 1.5. The reason is that it is possible for the EKS index to violate the fixed-weight bounds to the true index; in the present example, it is apparent from **L** that the welfare of country 4 cannot exceed $e^{0.693} = 2$ times that of country 1. Yet, from **E**, the EKS index suggests that the upper bound of the welfare of country 4 compared with country 1 is $e^{0.702} = 2.02$ times. Thus the EKS index violates the bounds to the true index provided by the fixed-weight indexes, and therefore cannot be a true index.

A version of the EKS which can be used for (non-true) multilateral comparisons is constructed by taking the mean of the columns of **E** or, equivalently, the negative of the mean of the rows of **E**. This produces the following vector: [-0.511,0.510,-0.065,0.191,-0.126], and element i gives the EKS estimate of the welfare of country i relative to the sample mean.

3.4 Application to 1980 ICP Data

In this section, the Ideal Afriat Index is constructed using 1980 data on 18 goods and services for 60 countries from International Comparisons Project (ICP) (see Appendix A for further details of the data).

The test of HARP was not satisfied for all 60 countries, so an iterative procedure was used to find the maximum number of countries consistent with homothetic preferences. A subset of four countries satisfying HARP (Argentina, Austria, Botswana and the U.S.) was arbitrarily chosen and then countries were iteratively added to this subset if they satisfied HARP. This resulted in a total of 42 countries sharing common homothetic preferences; the Ideal Afriat Index ($\mathbf{a}^*$) and the bounds to this index ($\mathbf{a}^+$ and $\mathbf{a}^-$) are presented in Table 3.2. The country rankings based on $\mathbf{a}^*$ are also in Table 3.2.[9]

Using the Ideal Afriat Index, the U.S. is ranked first with a true welfare level approximately 4.3 ($= e^{1.456}$) times that of the sample average. The poorest country was Ethiopia, for which true welfare in 1980 was only approximately 12.9 percent of the sample average.

For comparative purposes, the EKS index was constructed for the 42 countries which satisfy HARP. The EKS index comparing country i with the sample mean is constructed as a geometric mean of the EKS indexes comparing country k with each of the N countries. This can be shown to be equal to (in logarithmic form):

$$\log Q_i^{EKS} = \frac{\frac{1}{N}\{\sum_{i=1}^{N} L_{ik} - \sum_{j=1}^{N} L_{kj}\}}{2}.$$

Thus, the EKS index comparing each country with the sample mean is simply a vector of overall means of the column and negative row means of $\mathbf{L}$. This is exactly the same method for calculating the Ideal Afriat Index, except the latter involves $\mathbf{M}$.

The rankings produced by the Ideal Afriat and EKS indexes are very similar (the Pearson's correlation coefficient is 0.999) but there are some differences. For example, the pairwise rankings of the following countries are changed: Argentina and Brazil, Belgium

[9] A problem with this procedure for finding the largest set of countries satisfying HARP is that the choice of the initial four countries will have an impact on the composition of the final set of countries (as will the order in which additional countries are tested). In the present example, countries were selected in alphabetical order and this is likely to be the reason why 25 of the 30 countries in the first half of Table 3.2 satisfy HARP, compared with only 17 of the 30 countries in the second half of the table. In Chapter 4 this issue is addressed.

Table 3.2: Multilateral indexes, 1980

	Ideal Afriat Index				EKS Index			Ideal Afriat Index				EKS Index	
	$\mathbf{a}^*$	$\mathbf{a}^-$	$\mathbf{a}^+$	rank	$\log Q_k^{EKS}$	rank		$\mathbf{a}^*$	$\mathbf{a}^-$	$\mathbf{a}^+$	rank	$\log Q_k^{EKS}$	rank
Argentina	0.382	0.354	0.410	19	0.414	18	Japan	0.948	0.916	0.980	11	0.960	11
Austria	1.156	1.131	1.180	8	1.174	8	Kenya	-1.448	-1.472	-1.423	38	-1.464	38
Belgium	1.262	1.244	1.279	5	1.272	6	Korea	-0.212	-0.296	-0.129	27	-0.240	27
Bolivia	-0.527	-0.554	-0.501	29	-0.537	29	Luxembourg	1.253	1.226	1.280	6	1.288	5
Botswana	-0.840	-0.863	-0.817	32	-0.848	32	Madagascar	-1.357	-1.390	-1.325	37	-1.379	37
Brazil	0.398	0.372	0.424	18	0.400	19	Malawi	-1.950	-1.999	-1.901	41	-2.010	41
Cameroon	-1.148	-1.171	-1.124	35	-1.145	35	Mali						
Canada	1.445	1.424	1.465	2	1.461	2	Morocco	-0.761	-0.786	-0.735	31	-0.744	31
Chile	0.278	0.240	0.315	20	0.297	20	Netherlands						
Colombia	0.093	0.069	0.116	22	0.095	22	Nigeria						
Costa Rica	0.245	0.227	0.262	21	0.221	21	Norway	1.090	1.063	1.118	9	1.093	9
Ivory coast	-0.895	-0.913	-0.878	33	-0.875	33	Pakistan						
Denmark	1.275	1.255	1.296	3	1.288	4	Panama	-0.076	-0.094	-0.057	24	-0.085	25
Dominican Rep.							Paraguay						
Ecuador	-0.196	-0.219	-0.173	26	-0.201	26	Peru	-0.084	-0.127	-0.040	25	-0.083	24
El Salvador	-0.452	-0.476	-0.428	28	-0.466	28	Philippines						
Ethiopia	-2.046	-2.100	-1.992	42	-2.093	42	Poland						
Finland	0.933	0.915	0.951	12	0.959	12	Portugal	0.465	0.420	0.510	17	0.499	15
France	1.221	1.200	1.243	7	1.244	7	Senegal	-1.160	-1.191	-1.128	36	-1.199	36
Germany FR	1.273	1.250	1.297	4	1.305	3	Spain						
Greece	0.574	0.556	0.591	13	0.574	13	Sri Lanka	-0.653	-0.725	-0.581	30	-0.676	30
Guatemala	-0.054	-0.076	-0.032	23	-0.068	23	U.R. Tanzania						
Honduras							Tunisia						
Hong Kong	1.000	0.930	1.070	10	1.009	10	U.K.						
Hungary	0.490	0.473	0.508	16	0.482	17	U.S.	1.456	1.438	1.473	1	1.494	1
India	-1.646	-1.679	-1.614	39	-1.618	39	Uruguay	0.540	0.518	0.563	14	0.550	14
Indonesia	-0.990	-1.036	-0.944	34	-1.011	34	Venezuela	0.527	0.476	0.578	15	0.485	16
Ireland							Yugoslavia						
Israel							Zambia	-1.810	-1.836	-1.784	40	-1.823	40
Italy							Zimbabwe						

and Luxembourg, Denmark and Germany and Panama and Peru. Only Portugal experiences a change in ranking of more than one place (it slips from 17th place when welfare is measured by $\mathbf{a}^*$ to 15th when ranked with EKS index). It is also apparent that the EKS index is slightly more dispersed than the Ideal Afriat Index. With the EKS index, the U.S. has a welfare level approximately 4.5 times the sample average, while true welfare for Ethiopia is only 12.3 percent of the average. The difference in dispersion of the welfare indexes is also indicated by the standard deviation which is 1.034 for the Ideal Afriat Index and 1.049 for the EKS index.

The high correlation between the Ideal Afriat Index and EKS index might suggest that in practice the choice between these indexes will have little significance. However, as Neary (2000) has noted, small differences between index numbers can have significant implications for policy issues when one allows for the cumulative effect over time. The U.S. Boskin Commission found a large impact on the social security bill associated with using differing methods for calculating the cost-of-living used in the indexing of pensions and other cash benefits. Neary (2000) simulated the impact of differing aggregation methods on aid flows using the 1980 ICP data. He assumed that each OECD country transferred 0.7 percent of real consumption expenditure (grossed up by population) and found that the Geary index implied a transfer which is over US\$0.3 billion greater than that implied by the EKS index.

It is of interest to see the extent to which the use of multilateral comparisons tightens the bounds to the true index. Figure 3.4 presents normalised true and fixed-weight bounds to the true welfare index comparing each country to the U.S. Each set of bounds is normalised relative to the log of the EKS index (Q_{iUS}^{EKS}), for reasons to be discussed below. Thus the normalised true bounds to the (log) welfare index comparing country i and the U.S. are $\{-M_{iUS} - \log Q_{iUS}^{EKS}, M_{USi} - \log Q_{iUS}^{EKS}\}$ and the normalised fixed-weight bounds are $\{-L_{iUS} - \log Q_{iUS}^{EKS}, L_{USi} - \log Q_{iUS}^{EKS}\}$.

It is apparent from Figure 3.4 that the elements of $\mathbf{M}$ provide significantly tighter bounds to the true welfare index. For example, the fixed-weight bounds suggest that in 1980, welfare in Ethiopia was between 1.8 and 27.9 percent that of the U.S. However, the minimum path matrix suggests that the true gap between Ethiopia and the U.S. must in fact lie between 2.9 and 3.1 percent.[10]

[10]Because of the normalisation, these numbers are not directly observable in Figure 3.4 but can be

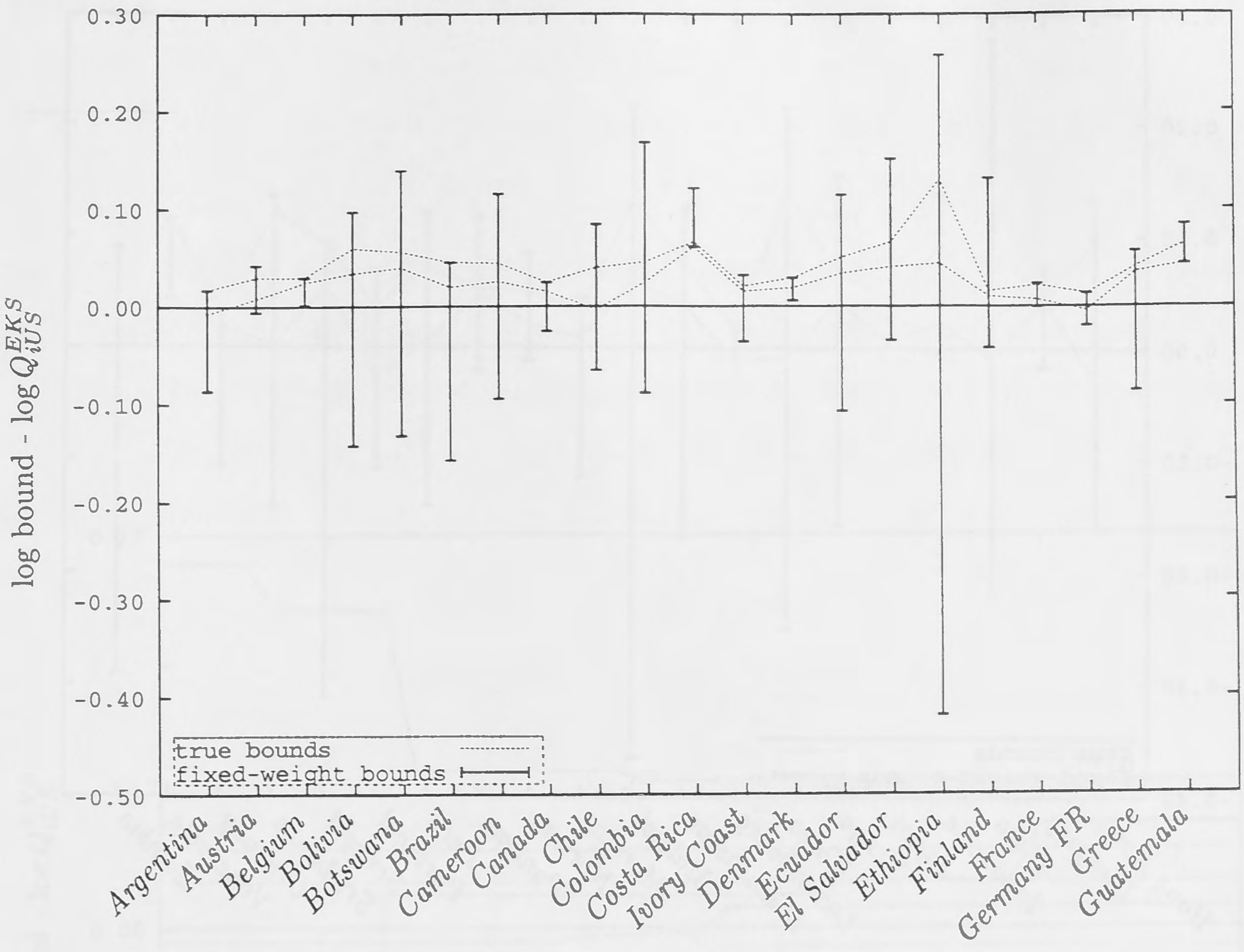

Figure 3.4: Normalised welfare bounds (U.S. base), 1980

Recall that a true index must lie within the P-L bounds (Definition 1.2). Because we have normalised the fixed-weight bounds on Q_{iUS}^{EKS} we are able to use Figure 3.4 to test whether the EKS index is a true index. For the several countries (Belgium, Costa Rica, Denmark, Guatemala, Japan and Venezuala), the normalised fixed weight bounds lie above the zero axis, thus indicating that the EKS index lies outside of the P-L bounds and for these data, therefore cannot constitute a true index.

Finally, it is of interest to see how the bounds to the true welfare index are tightened as the number of countries in the homothetic welfare comparison is increased. In Figure 3.5, the (normalised) true bounds to the true welfare index for three countries (where the U.S. is base) are plotted against the number of countries included in the welfare comparison. The three countries (Argentina, Austria and Botswana) are shown since they are present in the original subset of 4 countries which share HARP. Thus, the true bounds on the far left of the figure are calculated using the elements from $\mathbf{M}$ when $N = 4$, while the bounds

calculated using: $L_{USi} = -3.329$, $L_{iUS} = 4.007$, $M_{USi} = -3.456$ and $M_{iUS} = 3.543$.

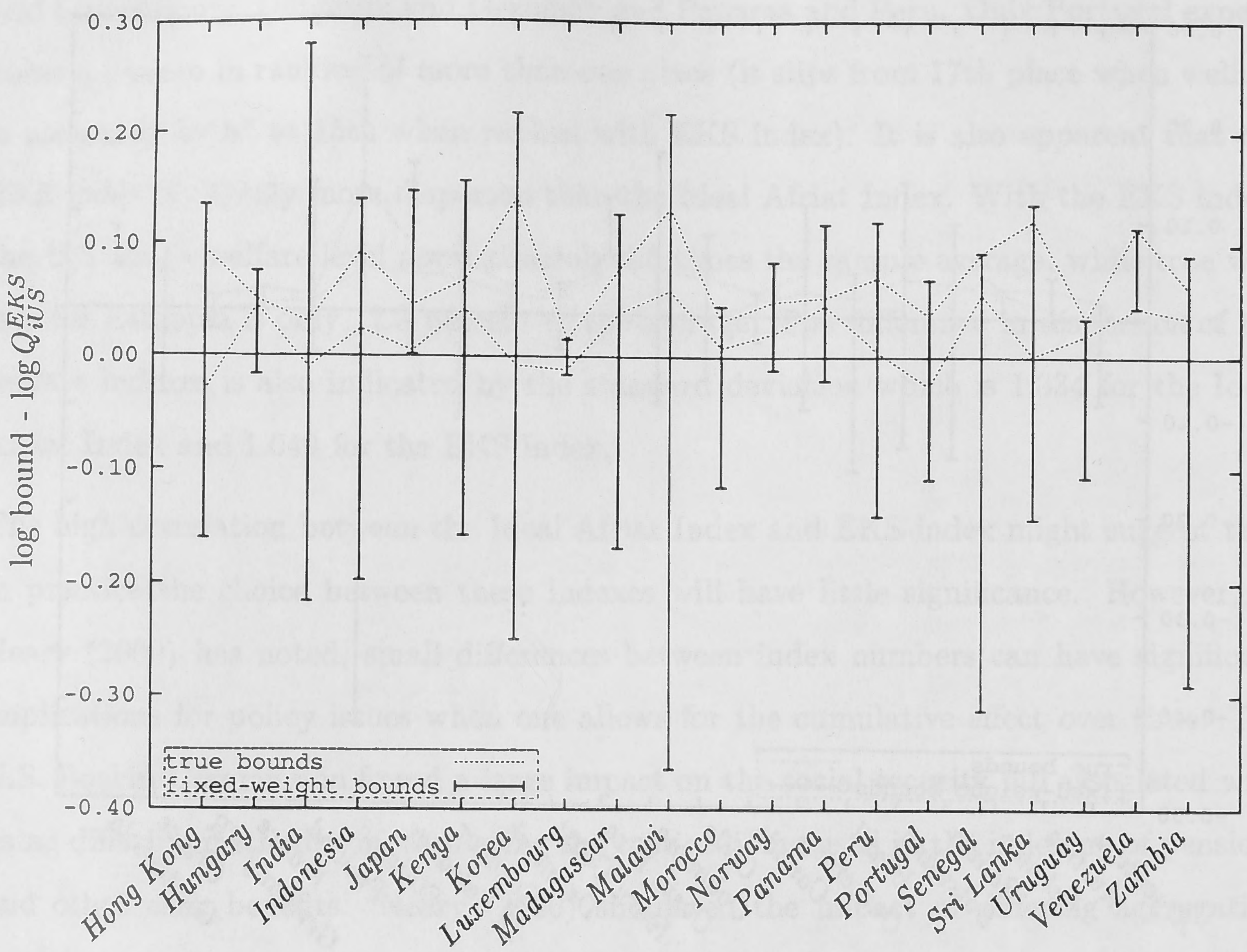

Figure 3.4: Normalised welfare bounds (U.S. base), 1980 - continued

on the far right of the figure are calculated from **M** when $N = 42$. As expected, adding more countries to the analysis progressively tightens the true bounds. With $N = 4$, the true welfare level of Botswana, for example, is between 8.4 and 11.0 percent of that of the U.S., while with $N = 42$ these bounds are significantly tightened so we can say that living standards in Botswana are between 10.0 and 10.2 percent of those in the U.S. (once again, because of the normalisation used these numbers are not directly observable in Figure 3.5).

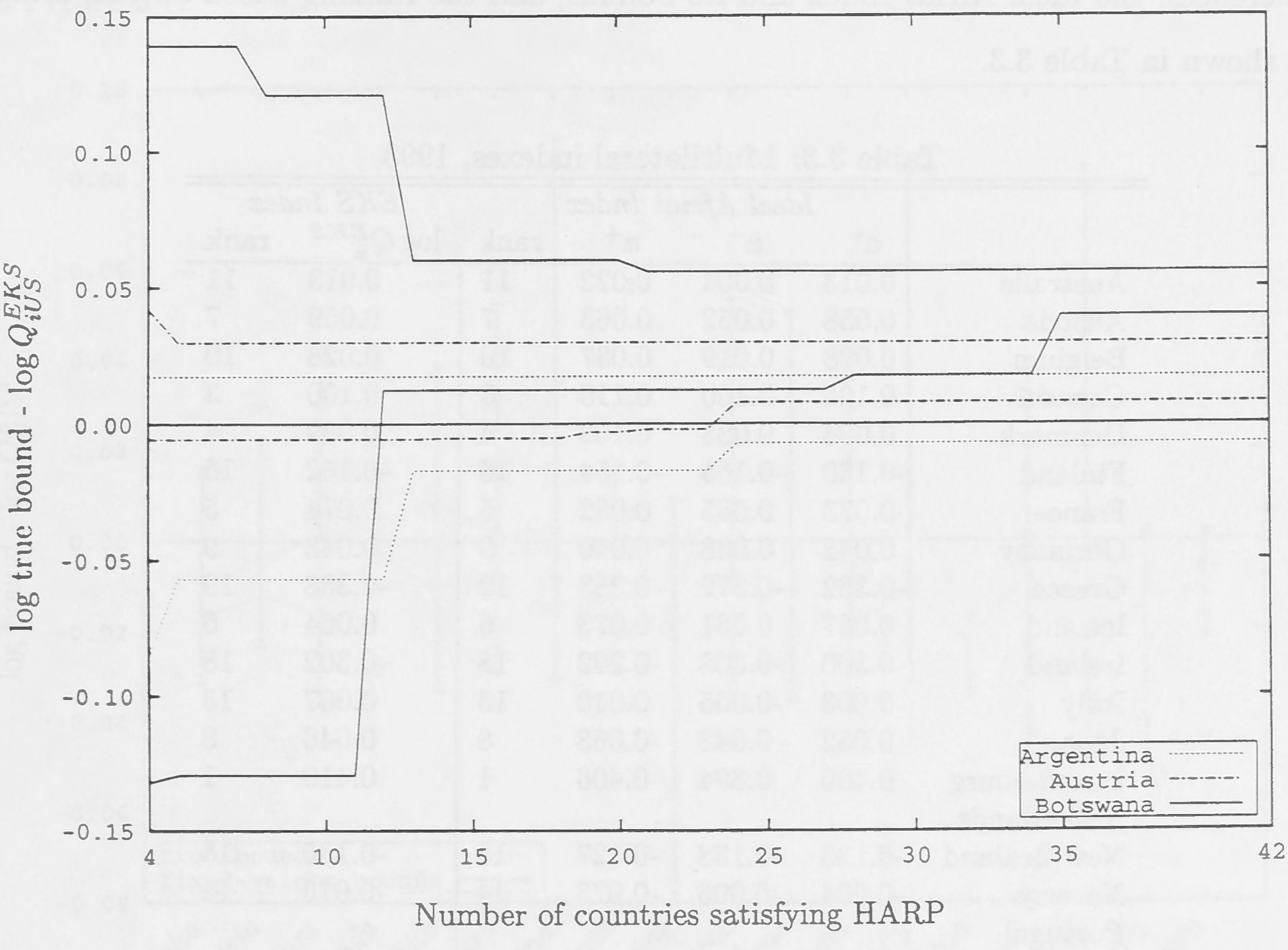

Figure 3.5: Normalised welfare bounds (U.S. base) as N increases, 1980

3.5 Application to 1993 ICP data

The Ideal Afriat Index is now constructed using 1993 ICP data on 19 goods and services for 24 OECD countries (see Appendix A for further details of the data).

As was found with the 1980 data, not all 24 countries share common homothetic preferences, so a subset of four countries satisfying HARP (Australia, Austria, Belgium and the U.S.) was arbitrarily chosen and the remaining countries were iteratively added to this subset if they satisfied HARP. This resulted in 19 countries sharing common homothetic preferences; the Ideal Afriat Index and its bounds, and the ranking based on this index are shown in Table 3.3.

Table 3.3: Multilateral indexes, 1993

	Ideal Afriat Index				*EKS Index*	
	$\mathbf{a}^*$	$\mathbf{a}^-$	$\mathbf{a}^+$	rank	$\log Q_k^{EKS}$	rank
Australia	0.013	0.004	0.022	11	0.013	11
Austria	0.058	0.052	0.063	7	0.059	7
Belgium	0.028	0.019	0.037	10	0.028	10
Canada	0.105	0.100	0.110	3	0.100	3
Denmark	0.094	0.085	0.103	4	0.089	4
Finland	-0.160	-0.165	-0.154	16	-0.162	16
France	0.073	0.065	0.082	5	0.074	5
Germany	0.043	0.038	0.049	9	0.043	9
Greece	-0.362	-0.372	-0.353	19	-0.358	19
Iceland	0.067	0.061	0.073	6	0.064	6
Ireland	-0.300	-0.308	-0.292	18	-0.302	18
Italy	0.003	-0.005	0.010	13	0.007	13
Japan	0.052	0.042	0.063	8	0.046	8
Luxembourg	0.400	0.394	0.406	1	0.410	1
Netherlands						
New Zealand	-0.133	-0.138	-0.127	15	-0.135	15
Norway	-0.084	-0.096	-0.073	14	-0.073	14
Portugal						
Spain	-0.267	-0.272	-0.262	17	-0.265	17
Sweden						
Switzerland						
Turkey						
U.K.	0.010	0.005	0.015	12	0.010	12
U.S.	0.359	0.353	0.365	2	0.352	2

Of the 19 OECD countries which share common homothetic preferences, Luxembourg is the richest with a level of real income approximately 50 percent above the sample average, and Greece is ranked last with a welfare level of 70 percent of the sample average. The EKS index (and implied ranking) is also shown in Table 3.3. The welfare rankings according to the EKS index are identical to those constructed using the Ideal

Afriat Index. In contrast to what was found with the 1980 data, the dispersion of welfare is approximately the same when measured by the EKS and Ideal Afriat indexes.

As with the 1980 data, we can graphically show how the bounds to the true welfare index are tightened by the use of multilateral comparisons. In Figure 3.6 the normalised bounds (with the U.S. as base) are shown. It is apparent that as was shown for the 1980 data, the EKS index constructed using the 1993 data is not a true index (since both the normalised fixed-weight bounds for Japan lie below the zero axis). The tightening of the welfare bounds for three countries as more countries are added to the analysis is shown in Figure 3.7.

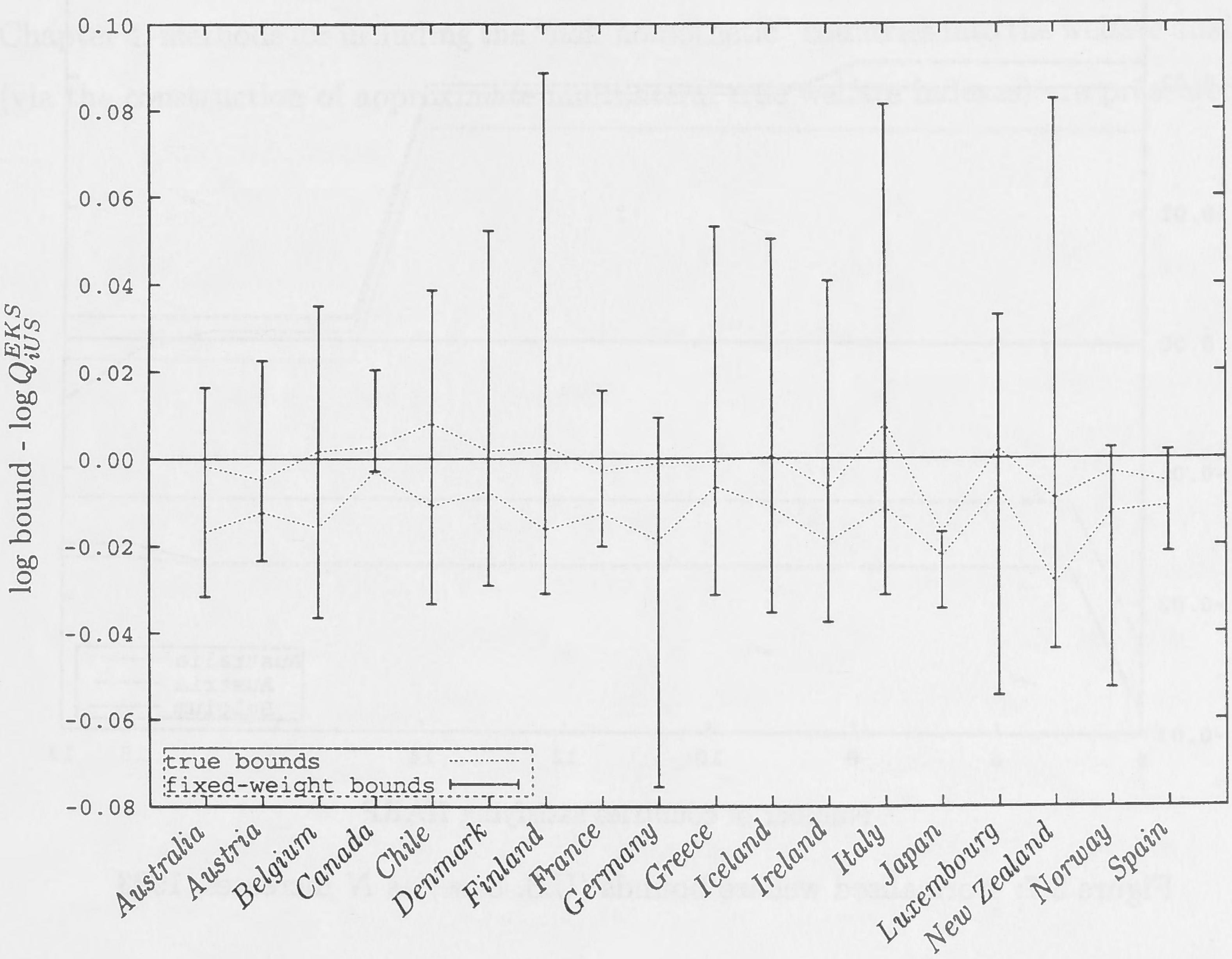

Figure 3.6: Normalised welfare bounds (U.S. base), 1993

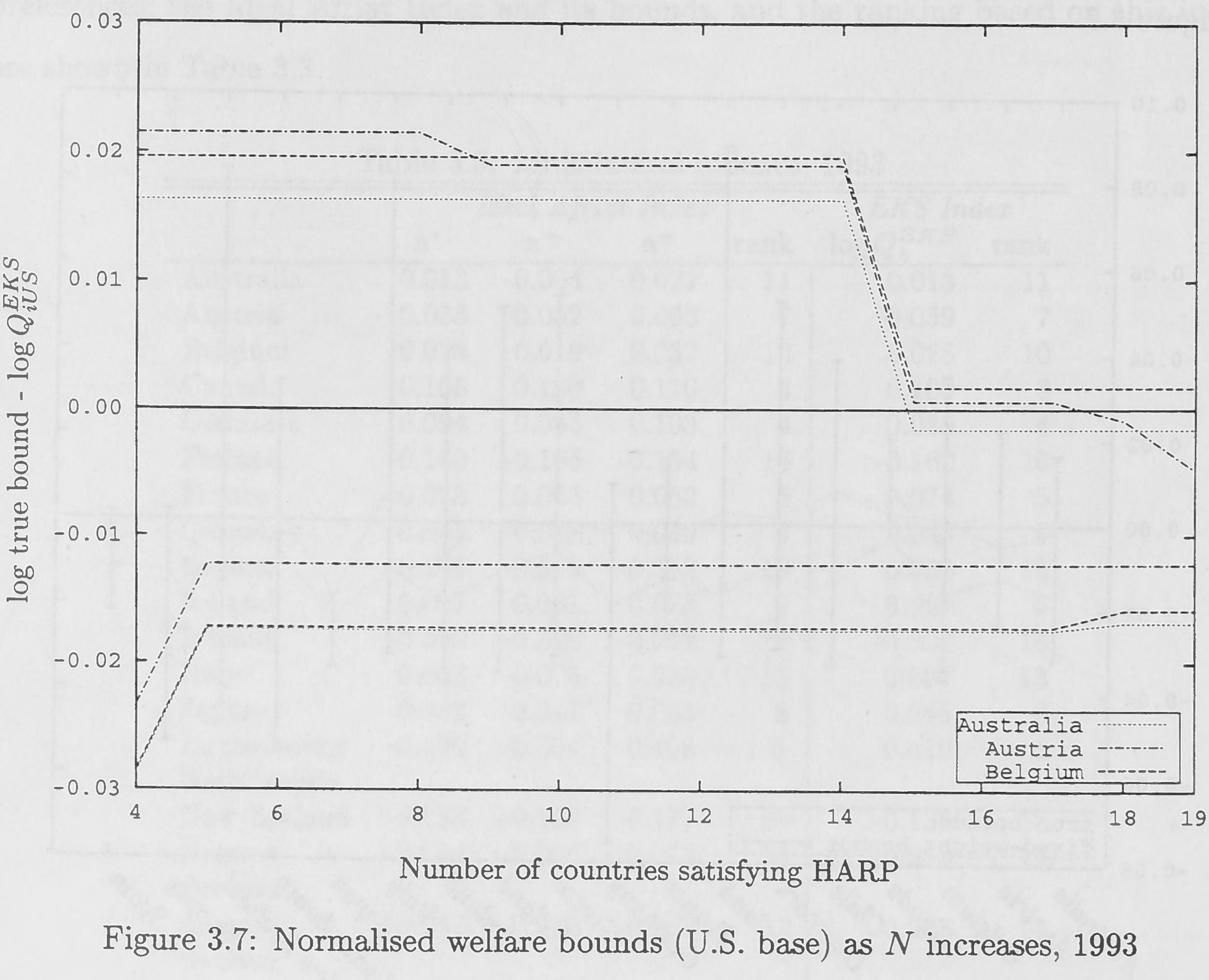

Figure 3.7: Normalised welfare bounds (U.S. base) as N increases, 1993

3.6 Conclusions

In this chapter, a review of the theory of the existence and construction of multilateral true indexes was presented. A particular multilateral true index, the Ideal Afriat Index of Dowrick and Quiggin (1997), has useful properties, and is the basis for the cross-country comparisons of welfare in the remainder of this thesis. The Ideal Afriat Index was constructed using 1980 and 1993 data from the ICP. It was found that of the 60 countries in the 1980 ICP data set, 42 share common homothetic preferences, while 19 of the 24 OECD countries in the 1993 data set satisfy HARP. The fact that utility consistent welfare comparisons cannot be made for all countries in the data sets is a problem and in Chapter 4, methods for including the "non-homothetic" countries into the welfare analysis (via the construction of approximate multilateral true welfare indexes) are presented.

Chapter 4

Approximate multilateral True Welfare Comparisons

4.1 Introduction

The revealed preference (RP) approach to welfare measurement reviewed in Chapters 2 and 3 is based on the notion of a utility maximising representative consumer, the existence of which is tested for via revealed preference relations. It was shown in Chapter 1 that homotheticity is necessary and sufficient for the existence of a unique welfare index, and for a finite data set, the most easily implemented test of homotheticity is the test of HARP (introduced in Proposition 3.1). For a given data set which satisfies HARP, there exists a multilateral true welfare index (defined in Definition 1.5) and the Ideal Afriat Index of Dowrick and Quiggin (1997) has been proposed as an appropriate multilateral true index for making utility consistent cross-country welfare comparisons.

One of the problems with the test of HARP (and the RP approach in general) is that it is non-stochastic - there is no allowance for the potential existence of errors in the vectors of quantities consumed. Such errors may be associated with measurement or with consumers not perfectly optimising their consumption choice, and may lead to the "all or nothing" HARP test rejecting common homothetic preferences for a particular data set. Varian (1990) has described RP tests as "sharp" in that the data either pass the tests *exactly* or else they don't and the hypothesis of utility maximisation (or of common preferences) is rejected: the tests do not allow for an "error term". The implications of this are apparent in Chapter 3, where HARP was satisfied for only 42 of the 60 countries in the 1980 data set, and 19 of the 24 countries in the 1993 data set.

There are several approaches which have been suggested for countering the overly sharp nature of RP tests. Manser and McDonald (1988), Patterson (1991) and Swofford and Whitney (1994) have devised various methods for perturbing data to ascertain the robustness of revealed preference tests of common preferences. In particular, Swofford and Whitney (1994) conduct tests of Type I error (where the null hypothesis of common preferences is true, but is rejected because of measurement error in the data) and Type II error (where measurement error leads to a false null hypothesis of common preferences being accepted). Varian (1985) used an alternative approach to find the minimal perturbation of the data required for the RP tests to be satisfied. However this approach is computationally intensive; testing data consisting of 18 observations on three factors of production of two electricity generation plants for consistency with the cost minimisation hypothesis involved a quadratic programming problem of 54 variables and around 200 constraints.

Banker and Maindiratta (1988) suggest finding the largest set of observations consistent with optimising behaviour (this was the approach employed in Chapter 4), although a problem is that the composition of the set will be dependent on the order of selection of observations for testing. Another obvious question is what to do about the observations which are not included in the largest set of consistent observations? In a cross-country context in particular, it may not be feasible or desirable to simply drop countries from the comparative analysis because they do not share common homothetic preferences with the other countries. Thus, for cross-country comparative work it is necessary to look at methods for making multilateral comparisons for all countries in the data set, even if they do not all share common homothetic preferences. There are two ways the RP approach can be modified to allow one to make utility-consistent multilateral comparisons amongst a set of countries which do not share common homothetic preferences.

The first approach is a method of imputing utility for those countries which fail the test of common homothetic preferences. This involves constructing the inner and outer homothetic envelope utility functions which bound all homothetic utility functions that rationalise the data (these are homothetic versions of the Afriat envelope functions reviewed in Section 2.4.4). Under the Afriat envelope approach, the set of countries is divided into two subsets: the set $\mathcal{H}$ which satisfy HARP, and the set $\mathcal{R}$ of countries which do not share common homothetic preferences with the countries in $\mathcal{H}$. For any county r in set $\mathcal{R}$, the method finds the utility of the representative consumer, whose preferences are homothetic and consistent with observations in the set $\mathcal{H}$, given they consumed the consumption bundle $\mathbf{q}^r$, even though the homothetic consumer would not have in fact chosen bundle $\mathbf{q}^r$ at prices $\mathbf{p}^r$. Thus the approach imputes the utility bounds for those countries which do not share common homothetic preferences.

The second approach to making utility comparisons for countries which do not share common homothetic preferences involves modifying the test of HARP itself. The Afriat critical cost efficiency index (Afriat (1967, 1972, 1987)) indicates how much each budget constraint has to be relaxed in order for the data to be consistent with utility maximisation. The Afriat efficiency index thus measures the overall "efficiency" of the consumer's choice behaviour and in the homothetic context, it is calculated as the smallest proportion of income that the consumer must waste if the data are to be shown to be consistent with homothetic preferences.

Varian (1990, 1993) used the Afriat efficiency index to define an approximate version of GARP for testing almost optimising behaviour. Afriat (1972, p.41) proposes a test of e *cyclical ratio consistency*, which is in fact the modified test of HARP employed in this chapter, as a test for almost homothetic optimising behaviour. In the present chapter, this previous research on testing data for consistency with almost optimising behaviour is extended to the construction of approximate multilateral true welfare indexes. In particular, an Ideal Afriat Index which incorporates specific levels consumer optimisation efficiency is proposed and constructed using ICP data.

There are two advantages to this approach. First, it enables utility-consistent comparisons to be extended to countries which do not satisfy the standard test of HARP. Second, the use of the Afriat efficiency index in the construction of multilateral true indexes allows one to evaluate the impact of the presence of optimising error on the welfare ranking between countries. In particular, if the pairwise ranking between two countries changes under differing assumptions regarding the level of consumer optimisation error (up to an appropriate maximum), then it can be argued that two countries should really be ranked as equivalent.

In this chapter, both the Afriat envelope and Afriat efficiency index approaches to making approximate multilateral comparisons are reviewed and implemented in an example. Of the two approaches, the Afriat efficiency index method is preferred since, unlike the Afriat envelope approach, it directly tackles the existence of errors in quantities which may be leading to the failure of HARP. Under the Afriat envelope approach, we calculate what the bounds to the utility of country r would be if that country had the same preferences as the countries in $\mathcal{H}$. There is no attempt to model the errors (either in $\mathbf{q}^r$ or in the consumption bundles of the countries in $\mathcal{H}$) that may be leading to the failure of HARP; rather, it is implicitly assumed that the data are accurate and country r simply does not share common homothetic preferences.

In contrast, the Afriat efficiency index approach to approximate multilateral comparisons directly models the existence of errors in quantities consumed and calculates the level of consumer optimising error that is required for the data to be consistent with approximate homothetic utility maximisation. For this reason, the Afriat efficiency index approach is preferred to the Afriat envelope approach and it is implemented here with 1980 and 1993 ICP data.

It is found that while 42 of the 60 countries in the 1980 data set share common homothetic preference using the standard test of HARP, 59 countries satisfy HARP when 4.2 percent consumer optimisation error is allowed for. Further, it is shown that two pairs of countries (Denmark and Germany, and Belgium and Luxembourg) should perhaps be ranked equivalent, since their pairwise rankings change with only moderate levels of consumer optimisation error. While 19 of the 24 OECD countries in the 1993 data set satisfy HARP with zero optimisation error, all satisfy HARP when 2 percent optimisation error is allowed for. Further, it is found that the rankings of Italy, Australia and the U.K. are indeterminate and thus these countries should perhaps be ranked as equivalent.

The structure of this chapter is as follows. In Section 4.2, the Afriat envelope approach to imputing utility bounds is reviewed and implemented using a two-good six-country example. The Afriat efficiency index approach to approximate multilateral welfare comparisons is presented in Section 4.3 and illustrated using the example data. In Section 4.4, the Afriat efficiency index approach is applied to the 1980 ICP data, and an application to the 1993 ICP data is in Section 4.5. In section 4.6, the conclusions of this chapter are presented.

4.2 Approximate multilateral Comparisons Using Afriat Envelope Functions

In this section, the Afriat envelope approach to approximate multilateral comparisons is described and implemented in an example. A given set of countries can be divided into two subsets: the set $\mathcal{H}$ which satisfy HARP, and the set $\mathcal{R}$ which do not share common homothetic preferences with the countries in $\mathcal{H}$.

4.2.1 Conditional Afriat homothetic envelope functions

The Afriat numbers calculated as part of the mathematical programming test for common homothetic preferences (equation 3.5 in Section 3.2.2) are not unique; underlying this is the fact that for any set of homothetic demand data, there will be a whole family of homothetic utility functions that are consistent with the data and the utility maximisation hypothesis. The results on constructing Afriat envelope functions in Section 2.4.4 are now adapted to show that a family of homothetic utility functions may be bounded by particular representations of the utility function, known as the *inner* and *outer homothetic envelope functions* for the class of compatible utility functions associated with any particular solution to the homothetic Afriat inequalities (equation 3.3 in Proposition 3.1).

The inner homothetic envelope function is a monotonic concave polytope function and is defined for a given set of A satisfying the Afriat inequalities (3.3) as:

$$a_I(\mathbf{q}, \boldsymbol{A}) = \max_{\boldsymbol{\theta}} \Big[\sum_{i\in\mathcal{H}} A^i\theta^i : \sum_{i\in\mathcal{H}} \mathbf{q}^i\theta^i \leq \mathbf{q}, \sum_{i\in\mathcal{H}} \theta^i = 1, \ \theta^i \geq 0, \ i \in \mathcal{H}\Big]. \tag{4.1}$$

For a given set of A satisfying the Afriat inequalities (3.3) the outer envelope function is a monotonic concave polyhedral function and is defined:

$$a_O(\mathbf{q}, \boldsymbol{A}) = \min_i [A^i\mathbf{v}^i.\mathbf{q} : \ i \in \mathcal{H}]. \tag{4.2}$$

From Proposition 3.1 in Section 2.4.4, the following can be stated:

Lemma 4.1. *$a_I(\mathbf{q}, \boldsymbol{A})$ and $a_O(\mathbf{q}, \boldsymbol{A})$ are representations of consumer preferences which rationalise the demand data.*

Adapting Result 3.3 in Section 3.2.1 to the homothetic case, we can state the following.

Result 4.1. *For a given set of* $\boldsymbol{A}$ *satisfying* (3.3), *the functions* $a_I(\mathbf{q}, \boldsymbol{A})$ *and* $a_O(\mathbf{q}, \boldsymbol{A})$ *provide, respectively, the inner-bound and outer-bound homothetic representation of consumer preferences:*

$$a_I(\mathbf{q}, \boldsymbol{A}) \leq a(\mathbf{q}) \leq a_O(\mathbf{q}, \boldsymbol{A}),$$

where $a(\mathbf{q})$ *is any concave, monotonic, continuous, non-satiated homothetic utility function that rationalises the data and satisfies* $a(\mathbf{q}^i) = A^i$ *for all* $i \in \mathcal{H}$.

Result 4.1 shows that $a_I(\mathbf{q}, \boldsymbol{A})$ and $a_O(\mathbf{q}, \boldsymbol{A})$ provide the widest bounds to all possible concave and non-satiated homothetic utility representations of $u(\mathbf{q})$. However, it should be emphasised that these bounds are *conditional* on a particular set of A satisfying the Afriat inequalities (3.3).

4.2.2 Unconditional Afriat homothetic envelope functions

The non-uniqueness of the solution to the Afriat inequalities (3.3) implies that the conditional inner and outer homothetic envelope functions presented above will similarly not be unique. However, following Chavas and Cox (1997), we can calculate *unconditional* bounds for $a(\mathbf{q})$, i.e. bounds that do not depend on the values taken by A.

The concave function $a(\mathbf{q})$ is only defined up to a positive linear transformation and hence, without loss of generality, the Afriat inequalities (3.3) can alternatively be written as:

$$A^i/A^j \leq \mathbf{v}^j.\mathbf{q}^i, \quad i, j \in \mathcal{H}$$

$$A^b = 1. \tag{4.3}$$

where A^b is the utility level for some base observation $b \in \mathcal{H}$.

We are now able to define the *unconditional* homothetic inner bound at point $\mathbf{q}$:

$$a_{UI}(\mathbf{q}) = \min_{\boldsymbol{A}}[a_I(\mathbf{q}, \boldsymbol{A}) : \text{ Eq. (4.3)}]. \tag{4.4}$$

The unconditional homothetic outer bound at point $\mathbf{q}$ is similarly defined as:

$$a_{UO}(\mathbf{q}) = \max_{\boldsymbol{A}}[a_O(\mathbf{q}, \boldsymbol{A}) : \text{ Eq. (4.3)}]. \tag{4.5}$$

Result 4.2. *Given (4.3), the functions $a_{UI}(\mathbf{q})$ in (4.4) and $a_{UO}(\mathbf{q})$ in (4.5) provide, respectively, the unconditional inner- and outer-bound homothetic representations of consumer preferences at point $\mathbf{q}$:*

$$a_{UI}(\mathbf{q}) \leq a(\mathbf{q}) \leq a_{UO}(\mathbf{q}).$$

Result 4.2 shows that the functions $a_{UI}(\mathbf{q})$ and $a_{UO}(\mathbf{q})$ provide the widest possible bounds on all possible concave and non-satiated homothetic utility representations of $u(\mathbf{q})$; because these bounds are constructed using 3rd party comparisons, they will generally be tighter than the fixed-weight (P-L) bounds (and cannot be wider than these bounds).

Chavas and Cox (1997) presented mathematical programming methods for calculating $w_{UI}(\mathbf{q})$ and $w_{UO}(\mathbf{q})$ (the unconditional bounds to consumer preferences in the general case) and constructing the implied money metric utility bounds (see Section 2.4.4 for details). However, with homotheticity, the money metric utility bounds are provided directly from $a_{UI}(\mathbf{q})$ and $a_{UO}(\mathbf{q})$ that is, it is not necessary to take the additional step of constructing inner and outer expenditure functions. Further, the functions $a_{UI}(\mathbf{q})$ and $a_{UO}(\mathbf{q})$ are easily constructed using the elements of the minimum path matrix $\mathbf{M}$.

Proposition 4.1. *The unconditional homothetic inner and outer bound envelope functions are calculated as, respectively:*

$$a_{UI}(\mathbf{q}) = \max_{\boldsymbol{\theta}} \Big[\sum_{i\in\mathcal{H}} \exp(-M_{ib})\theta^i : \sum_{i\in\mathcal{H}} \mathbf{q}^i\theta^i \leq \mathbf{q},\ \sum_{i\in\mathcal{H}} \theta^i = 1,\ \ \theta^i \geq 0,\ \ i \in \mathcal{H}\Big] \tag{4.6}$$

$$a_{UO}(\mathbf{q}) = \min_i[\exp(M_{bi})\mathbf{v}^i.\mathbf{q} :\ \ i \in \mathcal{H}], \tag{4.7}$$

where M_{ij} is the ijth element of the minimum path matrix, $\mathbf{M}$, $\mathcal{H}$ is the set of countries which share common homothetic preferences, and $A^b = 1$.

Proof: For a data set which is consistent with homotheticity, the definition of $\mathbf{M}$ (see Section 3.3) implies that $\exp(-M_{ib})$ is a tight lower bound to A^i when $A^b = 1$. Thus the numbers $\exp(-M_{.b})$ form a vector of tight lower bounds to the utility of the homothetic countries when $A^b = 1$, and the unconditional homothetic inner bound envelope function (4.5) follows.

Similarly, the number $\exp(M_{bi})$ gives the upper bound to A^i consistent with the data being rationalised by a homothetic utility function and $A^b = 1$. Thus the unconditional homothetic outer bound envelope function can be constructed using the bth row of $\mathbf{M}$ as (4.7).□

For $\mathbf{q} = \mathbf{q}^i$, the value of the unconditional homothetic inner bound envelope function will be $a_{UI}(\mathbf{q}^i) = \exp(-M_{ib})$. Similarly, it is apparent that since $\mathbf{v}^i.\mathbf{q}^i = 1$, $a_{UO}(\mathbf{q}^i) = \exp(M_{bi})$, for all $i \in \mathcal{H}$. Further, as now shown, $a_{UI}(\mathbf{q})$ and $a_{UO}(\mathbf{q})$ can be used to construct lower- and upper-bounds to the utility of countries which are not in $i \in \mathcal{H}$ (and are thus not represented in the minimum path matrix).

4.2.3 Imputing utility bounds

The unconditional homothetic inner and outer envelope functions can be used to impute utility for countries which do not share common homothetic preferences. For any county r in $\mathcal{R}$, $a_{UI}(\mathbf{q}^r)$ and $a_{UO}(\mathbf{q}^r)$ provide the bounds to the utility that the representative consumer (whose preferences are homothetic and consistent with observations in the set $\mathcal{H}$) would enjoy if he or she consumed the bundle $\mathbf{q}^r$, even though the consumer in fact would not have chosen bundle $\mathbf{q}^r$ at prices $\mathbf{p}^r$.

Dowrick and Quiggin (1997) propose using a variant of (4.7) to impute utility for those countries which cannot be included in a homothetic representation of the data either because of missing price information or because of inconsistent preferences. However, Dowrick and Quiggin (1997) suggest using the Ideal Afriat Index in (4.7) rather than elements of $\mathbf{M}$, that is, they define the following utility function:

$$L(\mathbf{q}^r) = \min_{i \in \mathcal{H}} (a^{*i} + L_{ir}),$$

where r is a country in $\mathcal{R}$. This function returns a^{*i} for observations within $\mathcal{H}$ and an imputed utility (the minimum of the Laspeyres valuations) for those observations in $\mathcal{R}$.

However, this function will only define an upper bound to the Ideal Afriat Index for the countries not in $\mathcal{H}$. Dowrick, Dunlop, and Quiggin (1995) therefore propose the following counterpart utility function:

$$P(\mathbf{q}^r) = \max_{i \in \mathcal{H}} (a^{*i} - L_{ri}),$$

which again returns a^{*i} for observations within $\mathcal{H}$ and an imputed utility (the maximum of the Paasche valuations) for those observations not in $\mathcal{H}$. Dowrick, Dunlop, and Quiggin

(1995) further define the midpoint of these two functions:

$$A^*(\mathbf{q}^r) = [L(\mathbf{q}^r) + P(\mathbf{q}^r)]/2,$$

as an extension of the Ideal Afriat Index to include observations which do not fit the homothetic representation.

Note, however, that the function $P(\mathbf{q})$ is not the unconditional inner envelope function (4.6), and indeed, it is not apparent that it necessarily provides a representation of consumer preferences. Further, while by construction it must be that $a_{UI}(\mathbf{q}) \leq a_{UO}(\mathbf{q})$, as Dowrick, Dunlop, and Quiggin (1995) note, for a particular country r not in $\mathcal{H}$ it is not necessarily the case that $P(\mathbf{q}^r) \leq L(\mathbf{q}^r)$. If the aim of the utility imputation is to approximate homothetic behaviour for those countries not in $\mathcal{H}$ then the above finding is problematic.[1]

An alternative method of constructing $A^*(\mathbf{q})$ is to use the unconditional inner envelope function to construct the lower bound to the Ideal Afriat Index, that is, to construct:

$$P(\mathbf{q}^r) = \max_{\boldsymbol{\theta}} \,[\sum_{i\in\mathcal{H}} a^{*i}\theta^i : \;\sum_{i\in\mathcal{H}} \mathbf{q}^i\theta^i \leq \mathbf{q}^r, \;\sum_{i\in\mathcal{H}} \theta^i = 1, \;\theta^i \geq 0].$$

This ensures $P(\mathbf{q}^r) \leq L(\mathbf{q}^r)$, however there is still the remaining problem that the resulting index $A^*(\mathbf{q}^r)$ measures the utility of country r relative to the mean welfare level calculated *only over the homothetic countries.*

4.2.4 The Augmented Ideal Afriat Index

An alternative approach is to use the elements of the minimum path matrix (which, by definition, is constructed only over the countries in $\mathcal{H}$) to impute utility bounds using (4.6) and (4.7). This leads to the construction of an "augmented" minimum path matrix, $\mathbf{M}^a$, which includes all countries. The Augmented Ideal Afriat Index can then be constructed from $\mathbf{M}^a$, and it will measure utility relative to the welfare of all countries in the sample.

Definition 4.1. The Augmented Ideal Afriat Index: *For a given set of countries, let $\mathcal{H}$ represent the subset which satisfy HARP and $\mathcal{R}$ represent the subset which do not satisfy HARP. Using the minimum path matrix* $\mathbf{M}$ *(calculated for countries in*

[1]Another problem with this method is that it requires $\mathbf{p}^r$ and thus the method cannot be used if $\mathbf{p}^r$ has some missing elements.

$\mathcal{H}$), utility bounds can be imputed for the countries in $\mathcal{R}$ via (4.6) *and* (4.7)*. Let* $\mathbf{M}^a$ *represent the augmented minimum path matrix which contains the elements from* $\mathbf{M}$ *and the utility bounds imputed for the countries in $\mathcal{R}$. Let* $\mathbf{a}^{a+} \equiv (c^1, c^2, \ldots, c^N)$ *and* $\mathbf{a}^{a-} \equiv (-r^1, -r^2, \ldots, -r^N)$ *represent the vector of column means and the vector of negative row means, respectively, of* $\mathbf{M}^a$*. Define the Augmented Ideal Afriat Index as* $\mathbf{a}^{a*} \equiv (\mathbf{a}^{a+} + \mathbf{a}^{a-})/2$*, the vector of overall means.*

The properties and interpretation of the Augmented Ideal Afriat Index are analogous to that of the Ideal Afriat Index (see Proposition 3.4).

4.2.5 Utility approximation with the Afriat envelope approach - example

In Section 3.3.4, the minimum path matrix and Ideal Afriat Index were constructed for a two-good five-country example data set. The elements of $\mathbf{M}$ can be used to graph the unconditional homothetic inner and outer bound envelope functions using the approach outlined above. For this example, the unconditional envelope functions are constructed using country 3 as the base and hence $a_{UI}(\mathbf{q})$ is constructed using $\mathbf{M}_{.3}$ (Figures 4.1 and 4.2), while $a_{UO}(\mathbf{q})$ is constructed using $\mathbf{M}_{3.}$ and is shown in Figure 4.3. Together, the unconditional envelope functions bound all possible homothetic utility functions that are consistent with the data and normalised so that $A^3 = 1$. The unconditional inner and outer bounds to the indifference curve passing through $\mathbf{q}^3$ are shown in Figure 4.4; any (normalised) indifference curve for country 3 must pass within these bounds.

The Afriat envelope method of approximating utility can be illustrated by adding an extra country in the above example. As shown in Section 3.2.3, with country 6 included in the analysis (where $\mathbf{p}^6 = \{1/3, 1\}$ and $\mathbf{q}^6 = \{4.5, 2.5\}$), the data no longer satisfy HARP (this is because of a revealed preference violation between countries 2 and 6). However, utility consistent welfare comparisons may still be made for all six countries by using the elements of $\mathbf{M}$ (calculated for countries 1 to 5) to impute utility for country 6 via the functions $a_{UI}(\mathbf{q})$ and $a_{UO}(\mathbf{q})$. The approximate utility bounds for country 6 can

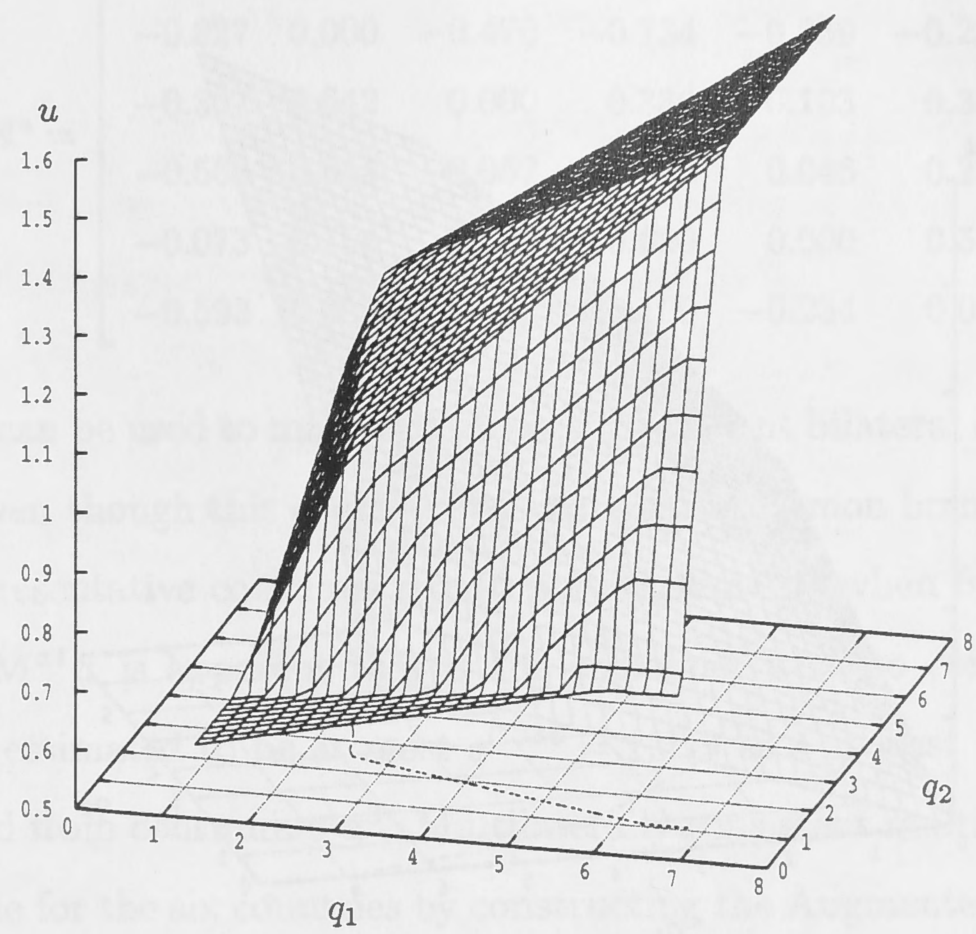

Figure 4.1: Unconditional homothetic inner envelope function, country 3 as base

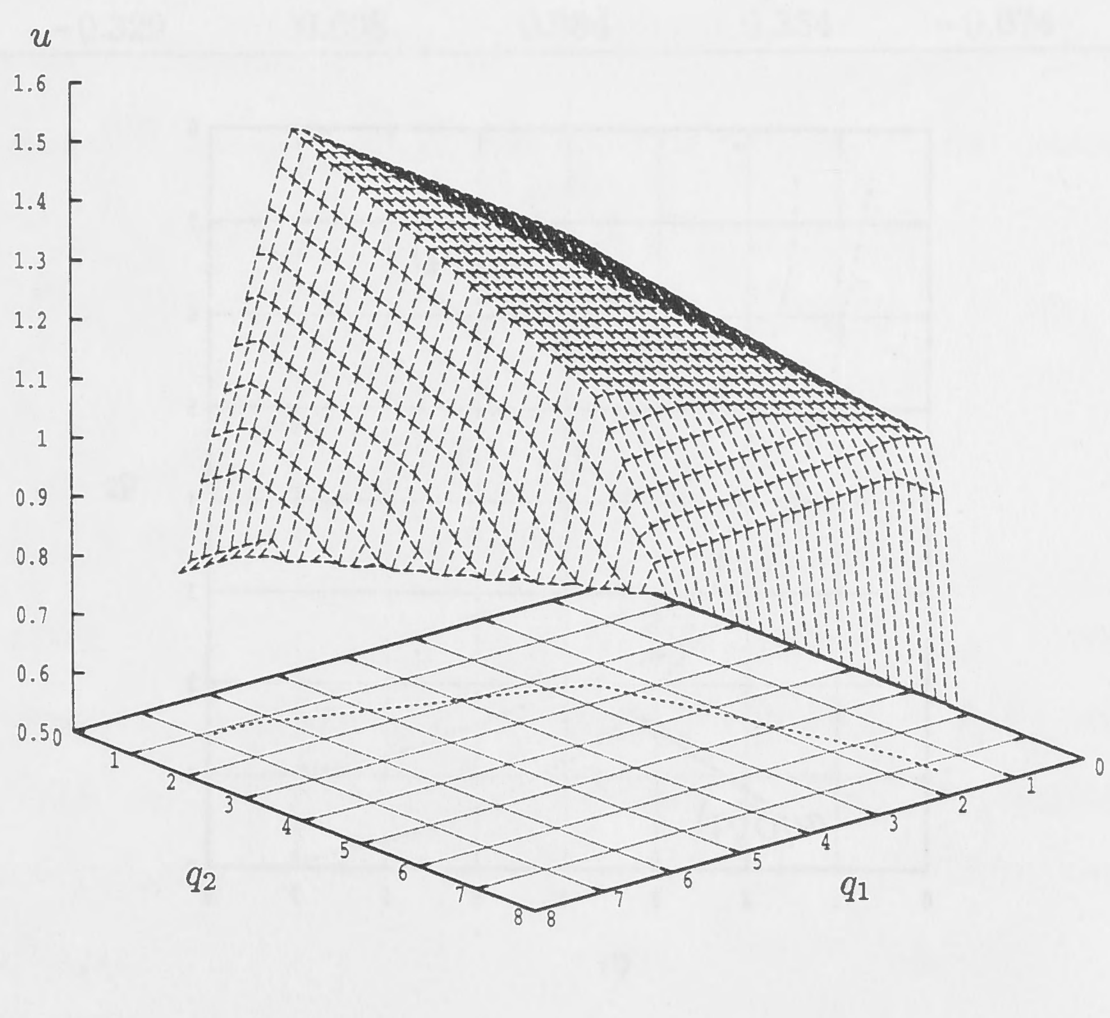

Figure 4.2: Unconditional homothetic inner envelope function (alternative view)

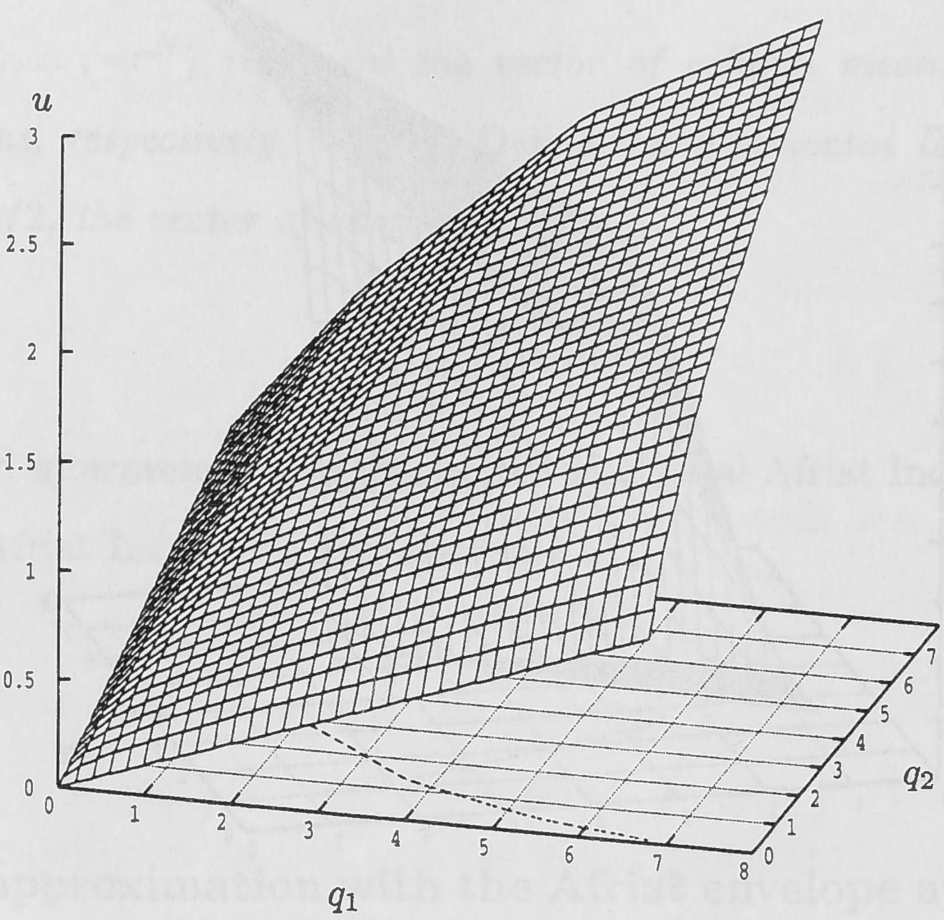

Figure 4.3: Unconditional homothetic outer envelope function, country 3 as base

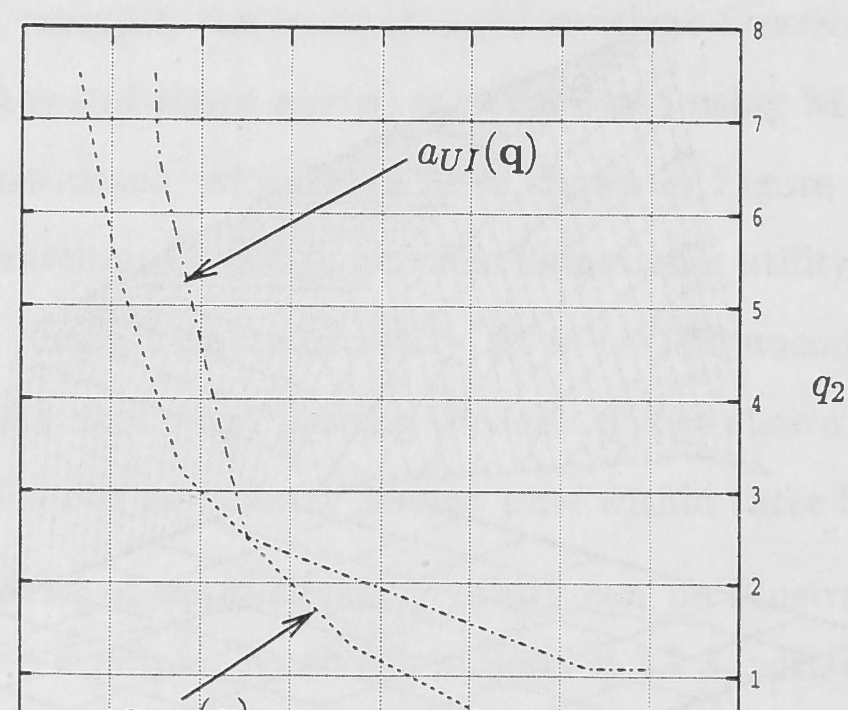

Figure 4.4: Unconditional homothetic inner and outer indifference curves for country 3

then be "appended" to $\mathbf{M}$ to form an augmented minimum path matrix:

$$\mathbf{M}^a = \begin{bmatrix} 0.000 & 1.153 & 0.511 & 0.693 & 0.614 & 0.847 \\ -0.827 & 0.000 & -0.470 & -0.134 & -0.539 & -0.234 \\ -0.357 & 0.642 & 0.000 & 0.336 & 0.103 & 0.336 \\ -0.568 & 0.585 & -0.057 & 0.000 & 0.046 & 0.279 \\ -0.073 & 0.754 & 0.284 & 0.620 & 0.000 & 0.520 \\ -0.593 & 0.305 & -0.236 & 0.100 & -0.234 & 0.000 \end{bmatrix}$$

The matrix $\mathbf{M}^a$ can be used to make multilateral consistent bilateral comparisons including country 6, even though this country does not share common homothetic preferences, and hence a representative consumer would not consume $\mathbf{q}^6$ when facing prices $\mathbf{p}^6$. For example, from $\mathbf{M}^a$ it is apparent that if the consumer were to consume $\mathbf{q}^6$, then the utility gained is estimated to be at most $e^{0.336} = 1.399$ and at least $e^{0.236} = 1.266$ times the utility gained from consuming $\mathbf{q}^3$. Multilaterally consistent multilateral comparisons can then be made for the six countries by constructing the Augmented Ideal Afriat Index (see Table 4.1).

Table 4.1: Augmented Ideal Afriat Index and bounds

	Country 1	Country 2	Country 3	Country 4	Country 5	Country 6
$\mathbf{a}^{a*}$	−0.447	0.497	−0.013	0.184	−0.221	0.256
$\mathbf{a}^{a-}$	−0.564	0.386	−0.110	0.014	−0.368	0.142
$\mathbf{a}^{a+}$	−0.329	0.608	0.084	0.354	−0.074	0.370

4.3 Approximate multilateral Comparisons Using the Afriat Efficiency Index

It has been shown that the Afriat envelope functions can be used to impute utility bounds for countries which do not share common homothetic preferences. Thus, it is possible to make multilaterally consistent welfare comparisons for all countries in a data set, even if HARP is not satisfied for the entire set of countries. While this may be useful in applied work, a potential limitation of the approach is that it does not directly address the reasons why a particular country may not share common homothetic preferences. In particular, there is an implicit assumption that there is no measurement or consumer optimisation error and if a country is not included in the set of homothetic countries then this is because that country truly does not share common homothetic preferences. Thus, the Afriat envelope approach to imputing utility can really be seen as an extension to the "all or nothing" RP approach, and does not address the presence of errors in the consumption vector which may lead the failure of HARP.

In reality, however, it is to be expected that the quantities used in multilateral comparisons are subject to measurement and consumer optimisation error. In the example of the previous section, the failure of country 6 to share common homothetic preferences could be due to the representative consumer in country 6 not successfully consuming the Marshallian demand bundle, or error in measuring exactly what the consumer purchased. For that matter, it could be the presence of errors in the consumption vector for country 2 which are leading to the revealed preference violation with country 6, yet the Afriat envelope approach to imputing utility does not allow for this.

An approach which does directly address the existence of measurement and consumer optimisation error involves the concept of the *Afriat critical cost efficiency index* of Afriat (1967, 1972, 1987), which shows how much each budget constraint has to be relaxed in order for the data to be consistent with utility maximisation. In the homothetic context, the Afriat efficiency index is calculated as the smallest proportion of income that the consumer must waste if the data are to be shown to be consistent with homothetic preferences. In constructing approximate multilateral welfare comparisons using the Afriat efficiency index, the researcher must decide how much consumer optimising error is considered acceptable; as suggested by Varian (1990), the "magic number" used in significance tests in econometrics, 5 percent, is probably a reasonable guide. Note that in using this approach, one is implicitly assuming that the only source of error is consumer optimisation error (and thus the presence of any measurement error will be treated as consumer optimisation error).

4.3.1 Approximate revealed preference tests

Varian (1990, 1993) used results from Afriat (1967, 1972) to define an approximate version of GARP for testing "almost optimising behaviour". In this sub section, these results are first reviewed then, following Afriat (1972), applied to the special case of homothetic preferences to define an approximate test of "almost optimising" homothetic behaviour.

General preferences

The revealed preference concepts introduced in Section 2.2.1 can be modified to take account of the presence of consumer optimisation error.

Definition 4.2. *Let e^i for $i \in N$ be numbers with $0 \leq e^i \leq 1$.*

1. $\mathbf{q}^i$ *is directly revealed preferred to* $\mathbf{q}$ *at efficiency level* e^i ($\mathbf{q}^i R^D(e^i)\mathbf{q}$) *if* $e^i\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}$;[2]

2. $\mathbf{q}^i$ *is revealed preferred to* $\mathbf{q}$ ($\mathbf{q}^i R(e^i)\mathbf{q}$) *at efficiency level* e^i *if* $e^i\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}^j$, $e^j\mathbf{p}^j.\mathbf{q}^j \geq \mathbf{p}^j.\mathbf{q}^k$, ..., $e^m\mathbf{p}^m.\mathbf{q}^m \geq \mathbf{p}^m.\mathbf{q}$ *for some sequence of observations* $(\mathbf{q}^i, \mathbf{q}^j, ..., \mathbf{q})$. *In this case the relation* $R(e^i)$ *is said to be the transitive closure of the relation* $R^D(e^i)$.

The parameter e^i is the *Afriat efficiency index* for observation i. It shows how much less the potential expenditure on a bundle $\mathbf{q}^j$ has to be before it will be revealed worse than the observed choice $\mathbf{q}^i$. Thus, if e^i is 0.9 and $\mathbf{q}^j$ would cost only 5 percent less than the cost of $\mathbf{q}^i$, we would not consider this a significant enough difference in cost to conclude that $\mathbf{q}^i$ is preferred by the consumer to $\mathbf{q}^j$. Therefore, we are allowing the consumer to have a "margin of error" of $(1 - e^i)$.

[2] Note that $R^D(1)$ is the standard direct revealed preference relation.

Given the above, we can now make a definition of GARP which incorporates consumer optimisation error:

Definition 4.3. *The data satisfy GARP(e^i) if* $\mathbf{q}^i R(e^i)\mathbf{q}^j$ *implies* $e^j\mathbf{p}^j.\mathbf{q}^j \leq \mathbf{p}^j.\mathbf{q}^i$.

If $e^i \equiv 1$ then we have the standard test of GARP, while if $e^i \equiv 0$ then the data will trivially satisfy the test. A measure of the overall efficiency of optimising behaviour in the data set can be found by finding how close the e^i must be to 1 for the data to satisfy GARP(e^i).

Afriat (1967) proposed using a *uniform* bound in the calculation of the efficiency index. That is, let $e^i = e^j$ for all $i, j \in N$ and then find the largest number e^* such that $\mathbf{q}^i R(e^*)\mathbf{q}^j$ implies $e^*\mathbf{p}^j.\mathbf{q}^j \leq \mathbf{p}^j.\mathbf{q}^i$.[3] The number e^* is the *Afriat critical cost efficiency index*, and it measures how much every budget constraint must be relaxed in order for the data to appear to be consistent with utility maximisation. The Afriat efficiency index indicates that the consumer is "wasting" a fraction $1 - e^*$ of income at each observation. If e^* is close to one, the consumer is wasting very little income, while if e^* is small then the consumer is wasting quite a lot of income. In this sense, e^* measures the overall "efficiency" of the consumer's choice behaviour.

Homothetic preferences

The test of *e cyclical ratio consistency* proposed by Afriat (1972, p.42) is an application of the above results to the special case of homothetic preferences and thus defines a test of approximate homothetic behaviour. Afriat's test modifies HARP to account for the presence of consumer optimisation error in the following way:

Definition 4.4. *The data satisfy HARP(e^i) if for all distinct choices of indexes* $(i, j, ..., m)$ *we have* $\left(\frac{\mathbf{v}^i.\mathbf{q}^j}{e^i}\right)\left(\frac{\mathbf{v}^j.\mathbf{q}^k}{e^j}\right)\ldots\left(\frac{\mathbf{v}^m.\mathbf{q}^i}{e^m}\right) \geq 1$, *or equivalently,* $(\mathbf{v}^i.\mathbf{q}^j)\ (\mathbf{v}^j.\mathbf{q}^k)\ldots(\mathbf{v}^m.\mathbf{q}^i) \geq e^ie^j\ldots e^m$.

If, as before, one is employing a uniform efficiency number e, the test of HARP(e) is easily implemented by constructing the minimum path matrix incorporating a particular

[3] In order to identify which observations are leading to revealed preference violations Varian (1993) suggested an approach for calculating observation-specific efficiency indexes i.e. $e^i \neq e^j$ for all $i, j \in N$. In the present context, where the goal is the construction of multilateral true indexes rather than the identification of observations failing HARP, it seems appropriate to use a uniform bound in the calculation of the efficiency index.

level of the Afriat efficiency index, which can be denoted $\mathbf{M}^e$. This matrix is constructed by applying Warshall's algorithm to a modified matrix of log Laspeyres indexes: $\{L^e_{ij}\} = (\mathbf{v}^i.\mathbf{q}^j/e)$. If all elements on the diagonal of $\mathbf{M}^e$ are non-negative, then the data satisfy HARP(e). As above, one can define a uniform efficiency bound e^* which is the largest number such that HARP(e^*) is satisfied. If e^* is close to 1 then the observed choice behaviour can be said to be efficient.[4] Alternatively, one can say that there only need to exist a small level of optimisation error for the data to be consistent with homothetic preferences. Note that the interpretation of e^* is slightly different to that used in the non-homothetic case. If HARP(e^*) is satisfied then we can say that the data satisfy homotheticity when the level of consumer optimisation error is $(1/e^* - 1)100$ percent. For example, HARP(0.909) implies an optimisation error level of $(1/0.909 - 1)100 \approx 10$ percent.

4.3.2 The Approximate Ideal Afriat Index

The Afriat efficiency index approach to conducting tests of approximate homotheticity can be extended to the construction of an approximate multilateral true welfare index.

Definition 4.5. THE APPROXIMATE IDEAL AFRIAT INDEX: *Let* $\mathbf{a}^+(e^*) \equiv (c^1, c^2, \ldots, c^N)$ *and* $\mathbf{a}^-(e^*) \equiv (-r^1, -r^2, \ldots, -r^N)$ *represent the vector of column means and the vector of negative row means, respectively, of the minimum path matrix incorporating the Afriat critical cost efficiency level* e^*, $\mathbf{M}(e^*)$. *Define the Approximate Ideal Afriat Index as* $\mathbf{a}^*(e^*) \equiv (\mathbf{a}^+(e^*) + \mathbf{a}^-(e^*))/2$, *the vector of overall means.*

The properties and interpretation of the Approximate Ideal Afriat Index are analogous to that of the Ideal Afriat Index (see Proposition 3.4).

4.3.3 Example 1

The approach for incorporating optimisation error into welfare comparisons can be illustrated using the simple three country example in Figure 4.5.

For these data, the matrix of log Laspeyres indexes and the minimum path matrix are

[4] Note that if $e^* = 0$, any set of data will be consistent with homotheticity.

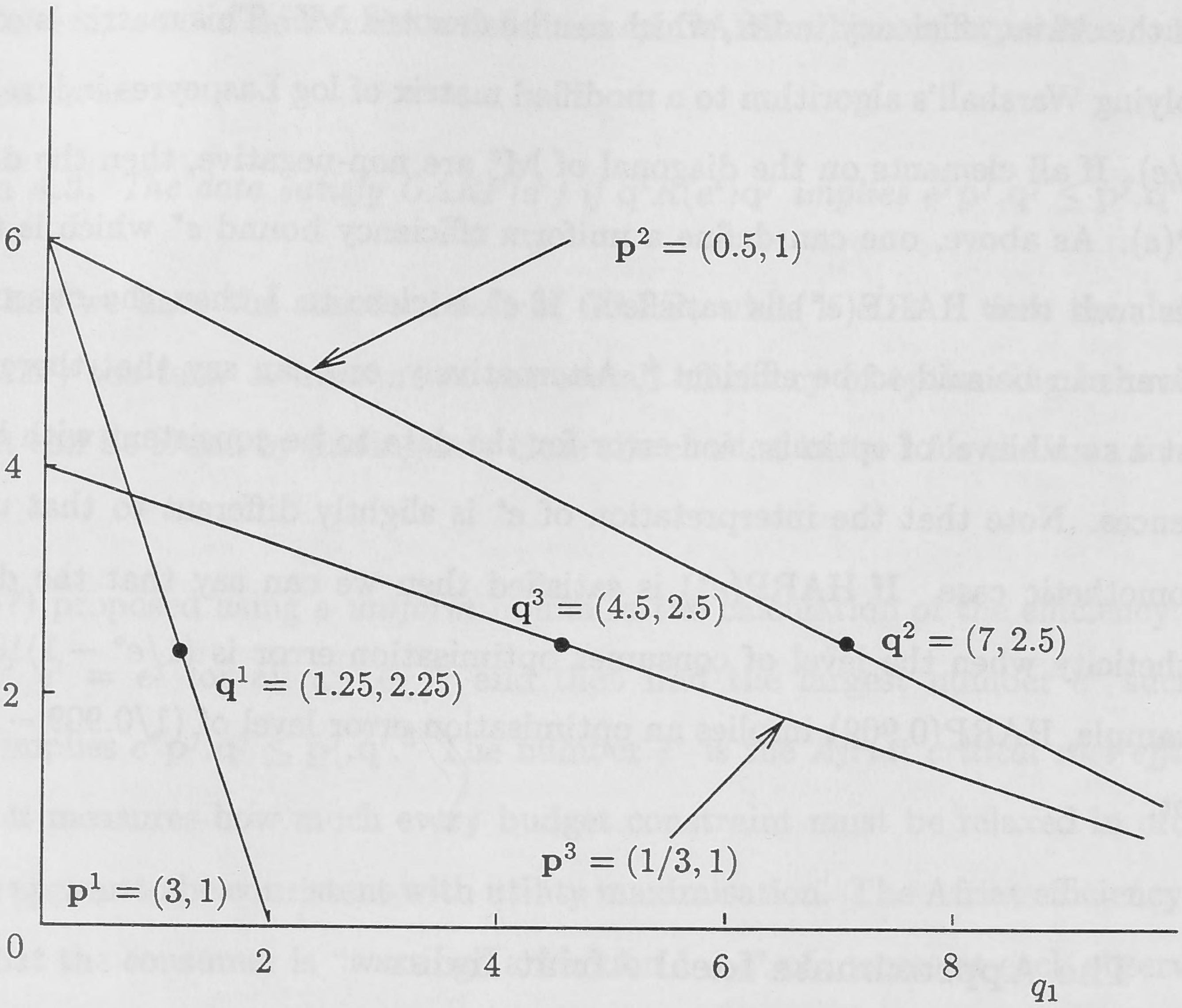

Figure 4.5: Three countries that do not share homothetic preferences

respectively:[5]

$$\mathbf{L} = \begin{bmatrix} 0.000 & 1.365 & 0.981 \\ -0.736 & 0.000 & -0.234 \\ -0.405 & 0.189 & 0.000 \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} 0.000 & 1.170 & 0.936 \\ -0.780 & -0.044 & -0.278 \\ -0.591 & 0.145 & -0.044 \end{bmatrix}$$

These data do not satisfy homotheticity and this is indicated by the fact that two of the diagonal elements of the minimum path matrix are negative (for example, $M_{33} = L_{32} + L_{23} = -0.044$). It is now shown how the presence of optimisation error can be incorporated into the data and thus how the data can be made to approximate homotheticity.

[5]Note that the minimum path algorithm in Varian (1983) would result in an **M** which is slightly different to that shown here. The algorithm used in the present example is changed for heuristic purposes (note, however that the different algorithms will not result in different minimum path matrices when HARP is satisfied, nor will they differ in whether a given data set satisfies HARP).

Assume that consumers have homothetic preferences but they do not optimise their consumption perfectly and hence are not able to achieve their Marshallian demand bundles. In particular, assume that the level of optimisation error is approximately 10 percent. In practical terms this means that if consumers optimised without error then they could have achieved the same level of utility spending 10 percent less of their budget, or alternatively, given the presence of optimisation error, consumers would have to spend 10 percent more of their budget to achieve the utility they would gain from consuming the Marshallian demand bundle.

From (3.3) in Proposition 3.1, the Laspeyres index $\mathbf{v}^i.\mathbf{q}^j$ gives an upper bound to A^j/A^i. With 10 percent optimisation error, all the Laspeyres indexes are increased by 10 percent, creating a new matrix of error-inclusive Laspeyres indexes $\{L_{ij}^{0.909}\} = \mathbf{v}^i.\mathbf{q}^j \times 1.1$. The interpretation is that for a given country j, the presence of optimising error means that $\mathbf{q}^j$ is not the Marshallian demand. With perfect optimisation, the consumption of $\mathbf{q}^j$ would give utility A^j, however the presence of optimising error (and homothetic preferences) means that the vector $\mathbf{q}^j \times 1.1$ would need to be consumed to achieve A^j. The Marshallian demand for consumer j is then found by finding the consumption bundle on j's budget constraint that will cost $\mathbf{p}^i.\mathbf{q}^j \times 1.1$.

For the above example, the cost of country 2's consumption vector at country 3's prices is $\mathbf{p}^3.\mathbf{q}^2 = 4.833$, but with 10 percent optimisation error the value of the Marshallian demand at country 3's prices is in fact $\mathbf{p}^3.\mathbf{q}^2 \times 1.1 = 5.316$. Thus, as shown in Figure 4.6, the Marshallian demand for country 2 can be shown to be $\mathbf{q}^{2M,3}$ where the notation indicates that $\mathbf{p}^3$ was used in the determination of this bundle. The Marshallian demand for country 3 using $\mathbf{p}^2$ can similarly be constructed (Figure 4.7). Note, however, that $\mathbf{q}^{2M,3}$ is not the only Marshallian demand for country 2 implied by the data - another demand bundle, $\mathbf{q}^{2M,1}$, can be found using country 1's prices as shown in Figure 4.8.

With optimisation error of 10 percent the data will now satisfy HARP. This can be seen by applying the minimum path algorithm to the modified matrix of log Laspeyres indexes:

$$\mathbf{L}^{0.909} = \begin{bmatrix} 0.000 & 1.461 & 1.076 \\ -0.640 & 0.000 & -0.138 \\ -0.310 & 0.285 & 0.000 \end{bmatrix}$$

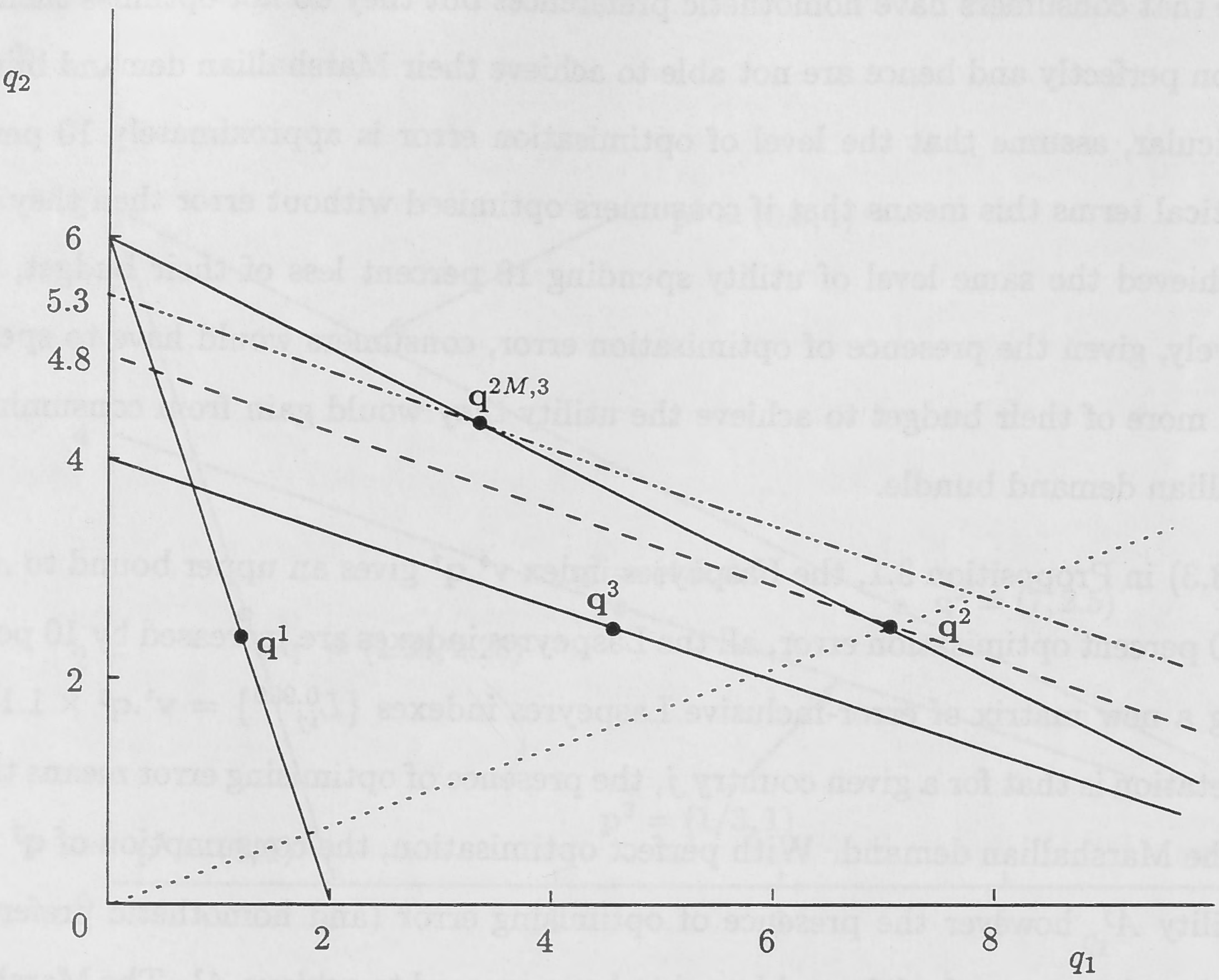

Figure 4.6: Finding country 2's Marshallian demand using $\mathbf{p}^3$

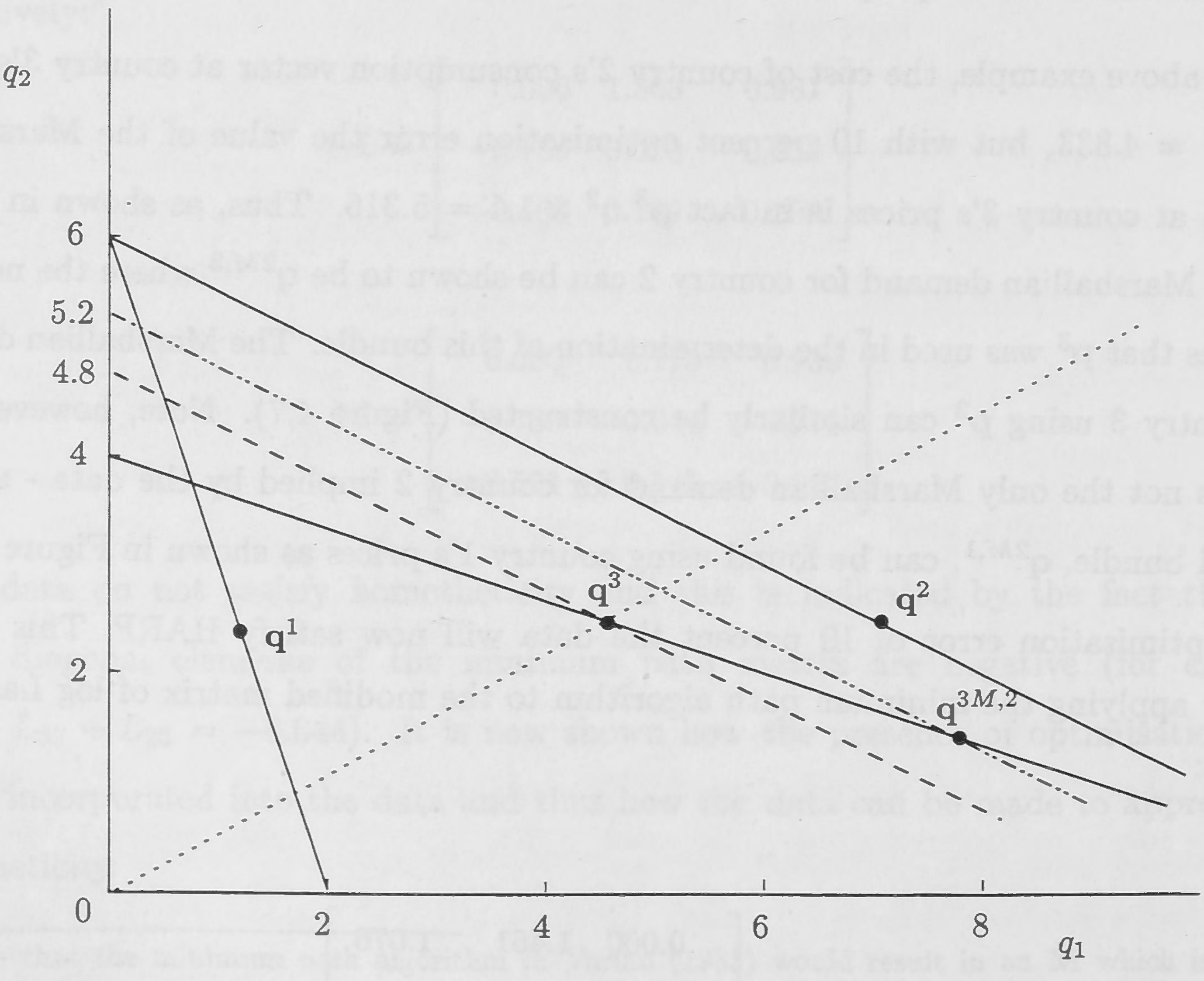

Figure 4.7: Finding country 3's Marshallian demand using $\mathbf{p}^2$

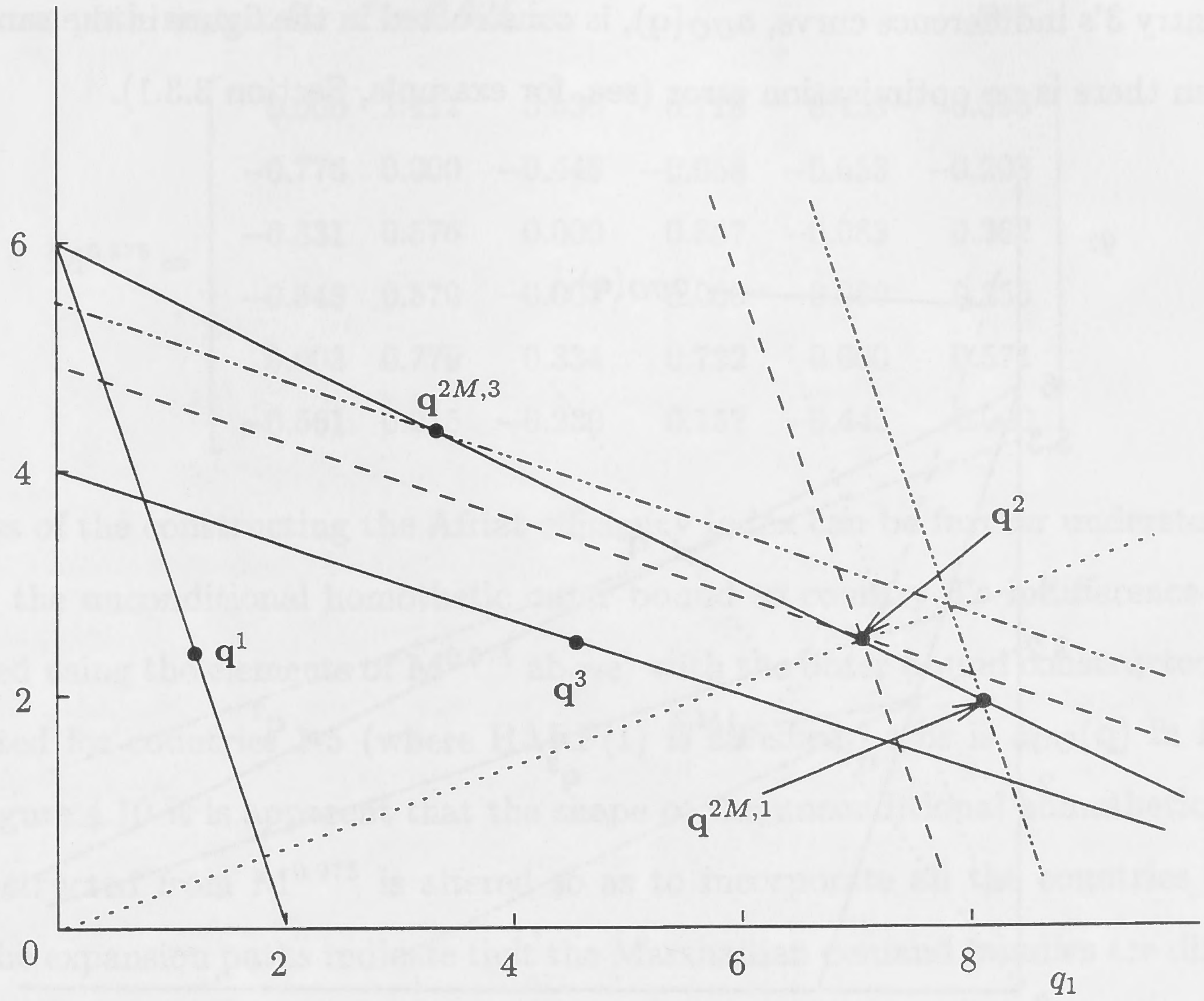

Figure 4.8: Country 2's Marshallian demands implied by $\mathbf{p}^1$ and $\mathbf{p}^3$

The minimum path matrix for these data is:

$$\mathbf{M}^{0.909} = \begin{bmatrix} 0.000 & 1.361 & 1.076 \\ -0.640 & 0.000 & -0.138 \\ -0.356 & 0.285 & 0.000 \end{bmatrix}$$

As there are no negatives on the diagonal, the data satisfy HARP (in particular, note that $L_{32}^{0.909} + L_{23}^{0.909} = 0.147$).

The elements of $\mathbf{M}^{0.909}$ provide the tightest possible bounds to the welfare index conditional on the presence of 10 percent optimisation error. For example, the element $M_{31}^{0.909}$, which is constructed as $\min\{L_{31}^{0.909}, L_{32}^{0.909} + L_{21}^{0.909}\} = -0.356$, gives the tightest upper bound estimate of A^1/A^3. Note that this upper bound is less tight than the fixed-weight bound with no optimisation error ($L_{31} = -0.405$) - but since the data do not satisfy homotheticity when there is no optimisation error, it is not clear in that case what the fixed-weight bound provides an upper bound to. The construction of this upper bound is illustrated in Figure 4.9. The Marshallian demand of country 1, calculated using $\mathbf{p}^2$ is denoted $\mathbf{q}^{1M,2}$ and is the bundle on country 1's constraint such that $\mathbf{p}^2.\mathbf{q}^{1M,2} = \mathbf{p}^2.\mathbf{q}^1 \times 1.1 = 3.163$. With both $\mathbf{q}^{1M,2}$ and $\mathbf{q}^{2M,3}$ determined, the outer bound to country 3's indifference curve, $a_{UO}(\mathbf{q})$, is constructed in the figure in the same manner as when there is no optimisation error (see, for example, Section 3.3.1).

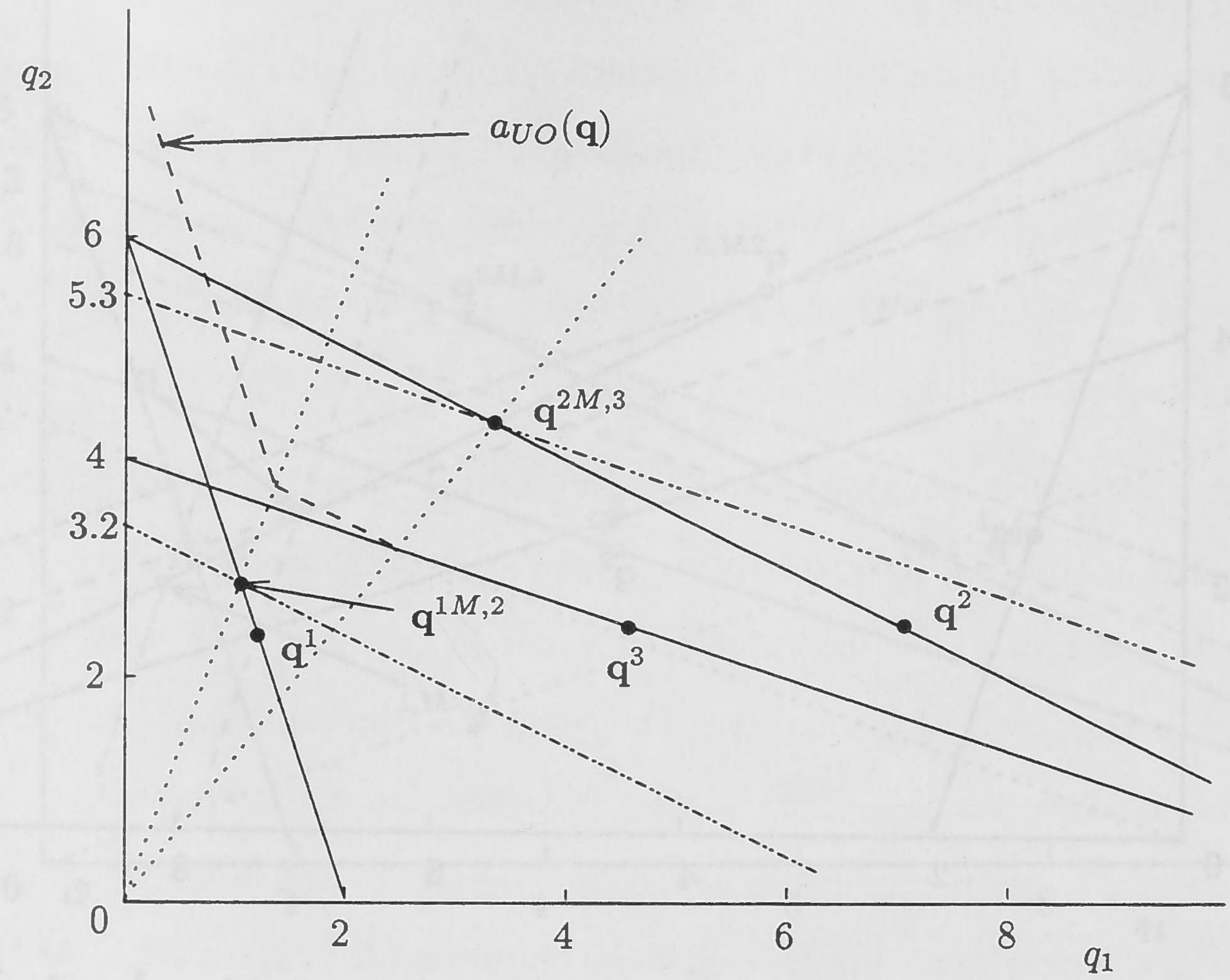

Figure 4.9: Constructing $a_{UO}(\mathbf{q})$ for country 3, 10 % optimisation error

4.3.4 Example 2

The Afriat efficiency index approach to approximate welfare comparisons is now illustrated using the six-country example data. These countries do not satisfy HARP and the question is: what is the minimum level of consumer optimisation error which leads to the data satisfying HARP? This question is answered by using a grid search procedure: e is initially set equal to 1, and then in each step is decreased by 0.005 and HARP(e) is tested.[6] The process was continued until HARP(e) is satisfied, and using this method, the data were found to satisfy HARP(0.975). That is, if the consumer is allowed to make optimisation error of 2.6 percent, then the data are consistent with homothetic preferences.

[6]Houtman and Maks (1987) suggested an alternative search procedure (which is more accurate than the one used in the present chapter). Start with $e = 1$ and test for HARP. If the data do not satisfy HARP, try $e = 1/2$. If $e = 1/2$ doesn't work, try $e = 1/4$. If the data *do* satisfy HARP with $e = 1/2$, try $e = 3/4$ and so on. After n revealed preference tests, one will be within $1/2^n$ of the actual efficiency index.

The minimum path matrix for $e^* = 0.975$ is:

$$\mathbf{M}^{0.975} = \begin{bmatrix} 0.000 & 1.112 & 0.536 & 0.718 & 0.453 & 0.898 \\ -0.776 & 0.000 & -0.445 & -0.058 & -0.653 & -0.208 \\ -0.331 & 0.576 & 0.000 & 0.387 & -0.083 & 0.362 \\ -0.543 & 0.570 & -0.007 & 0.000 & -0.089 & 0.355 \\ 0.003 & 0.779 & 0.334 & 0.722 & 0.000 & 0.571 \\ -0.561 & 0.215 & -0.230 & 0.157 & -0.445 & 0.000 \end{bmatrix}$$

The process of the constructing the Afriat efficiency index can be further understood by comparing the unconditional homothetic outer bound to country 3's indifference curve (constructed using the elements of $\mathbf{M}^{0.975}$ above) with the outer bound constructed from $\mathbf{M}$ calculated for countries 1-5 (where HARP(1) is satisfied - this is $a_{UO}(\mathbf{q})$ in Figure 4.4). In Figure 4.10 it is apparent that the shape of the unconditional homothetic outer bound constructed from $\mathbf{M}^{0.975}$ is altered so as to incorporate all the countries in the sample. The expansion paths indicate that the Marshallian demand bundles are different to the quantities actually consumed - the utility maximising bundles have "shifted" so that they are on the expansion paths that are consistent with HARP(0.975).

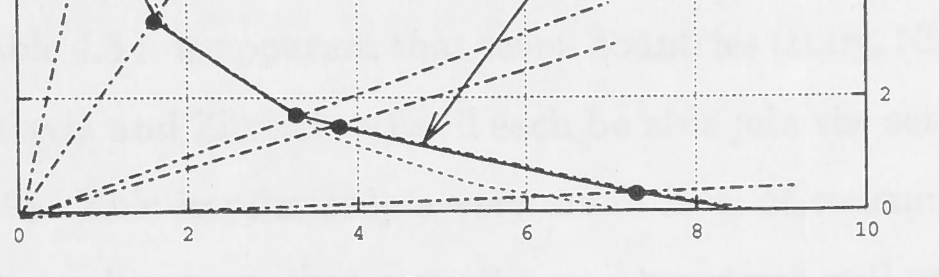

Figure 4.10: $a_{UO}(\mathbf{q})$ for country 3, constructed using HARP(0.975) and HARP(1)

Approximate multilateral welfare comparisons can be made using the Approximate Ideal Afriat Index calculated from $\mathbf{M}^{0.975}$ (Table 4.2).

Table 4.2: Approximate Ideal Afriat Index and bounds, $e^* = 0.975$

	Country 1	Country 2	Country 3	Country 4	Country 5	Country 6
$a^*(0.975)$	−0.494	0.449	−0.060	0.137	−0.269	0.237
$a^-(0.975)$	−0.620	0.357	−0.152	−0.048	−0.401	0.144
$a^+(0.975)$	−0.368	0.542	0.032	0.321	−0.136	0.330

4.4 Application to 1980 ICP Data

For reasons discussed above, the Afriat efficiency index approach is considered a preferable method for making approximate multilateral welfare comparisons. In this section, the Afriat efficiency index approach is applied to the 1980 ICP data.

As discussed in Section 3.4, the standard test of homothetic preferences, HARP(1), was only satisfied for 42 of the 60 countries in the 1980 data set (the Ideal Afriat Index and implied rankings are presented in columns 2 and 3 of Table 4.3 (these numbers are reproduced from Table 3.2). The set of 42 countries that satisfy HARP was found by first arbitrarily choosing a subset of four countries (Argentina, Austria, Botswana and the U.S.) and then iteratively adding countries to this subset if they satisfy HARP. A problem with this procedure for finding the largest set of countries satisfying HARP is that the choice of the initial four countries will influence the composition of the final set of countries (as will the order of selection of the remaining countries for testing). This problem is avoided by introducing consumer optimisation error into the test of HARP, since this will reduce the likelihood that a country is found to reject homotheticity because of the presence of another country in the set of homothetic countries.

As outlined above, a step procedure can be used to determine the maximum level of e^* which will allow a particular country to be consistent with homotheticity. This step procedure was applied to each of the 18 countries which do not satisfy HARP(1). From column 4 of Table 4.3 it is apparent that seven countries (Italy, Nigeria, Poland, Tanzania, Tunisia, Yugoslavia and Zimbabwe) will each be able join the set of homothetic countries when $e^* = 0.995$. This implies only a very small level of consumer optimisation error of 0.5 percent. Note, however, that not all seven countries will necessarily join the set of homothetic countries when e^* is set to 0.995 since there is likely to be inconsistencies of preferences amongst these countries which will not be resolved unless a smaller level of e^* (and thus higher level of optimisation error) is set. It is also apparent from column 4 that some countries require a large level of optimisation error before they are found to share common homothetic preferences; the Dominican Republic, for example requires $e^* = 0.96$.

Table 4.3: Approximate Ideal Afriat Index for different e^*, 1980

	$\mathbf{a}^*$	rank	e^*	$\mathbf{a}^*(0.99)$	rank	e^*	$\mathbf{a}^*(0.98)$	rank	e^*	$\mathbf{a}^*(0.97)$	rank	e^*	$\mathbf{a}^*(0.96)$	rank	e^*
Argentina	0.382	19		0.382	22		0.408	24		0.441	25		0.427	26	
Austria	1.156	8		1.154	9		1.166	9		1.203	9		1.205	9	
Belgium	1.262	5		1.270	4		1.275	4		1.312	5		1.306	6	
Bolivia	-0.527	29		-0.529	33		-0.521	38		-0.436	38		-0.479	41	
Botswana	-0.840	32		-0.840	36		-0.841	42		-0.830	43		-0.847	45	
Brazil	0.398	18		0.408	21		0.417	23		0.467	23		0.443	25	
Cameroon	-1.148	35		-1.153	39		-1.145	46		-1.114	47		-1.139	49	
Canada	1.445	2		1.442	2		1.447	2		1.491	2		1.486	2	
Chile	0.278	20		0.289	24		0.312	27		0.315	27		0.305	28	
Colombia	0.093	22		0.081	26		0.101	29		0.129	29		0.117	30	
Costa Rica	0.245	21		0.260	25		0.257	28		0.288	28		0.288	29	
Ivory coast	-0.895	33		-0.901	37		-0.897	43		-0.839	44		-0.856	46	
Denmark	1.275	3		1.267	5		1.274	5		1.317	4		1.318	4	
Dominican Rep.			0.96			0.96			0.96			0.96	-0.136	35	
Ecuador	-0.196	26		-0.189	30		-0.188	33		-0.161	34		-0.168	36	
El Salvador	-0.452	28		-0.457	32		-0.460	37		-0.443	39		-0.452	40	
Ethiopia	-2.046	42		-2.052	48		-2.059	55		-2.042	57		-2.053	59	
Finland	0.933	12		0.939	14		0.924	15		0.962	15		0.967	15	
France	1.221	7		1.229	7		1.240	7		1.280	7		1.281	7	
Germany FR	1.273	4		1.285	3		1.295	3		1.334	3		1.335	3	
Greece	0.574	13		0.563	16		0.568	18		0.633	18		0.603	19	
Guatemala	-0.054	23		-0.062	27		-0.063	30		-0.028	30		-0.046	31	
Honduras			0.985			0.965			0.965			0.96			0.94
Hong Kong	1.000	10		0.991	12		0.997	13		1.042	13		1.020	13	
Hungary	0.490	16		0.486	19		0.482	21		0.523	21		0.530	22	
India	-1.646	39		-1.643	44		-1.628	51		-1.581	52		-1.612	54	
Indonesia	-0.990	34		-0.974	38		-0.984	44		-0.916	45		-0.936	47	
Ireland			0.99	0.666	15		0.674	17		0.711	17		0.721	17	
Israel			0.975			0.975			0.96			0.96	0.693	18	
Italy			0.995	1.081	11		1.089	12		1.137	11		1.138	12	

Table 4.3: Approximate Ideal Afriat Index for different e^*, 1980 (cont.)

	$\mathbf{a}^*$	rank	e^*	$\mathbf{a}^*(0.99)$	rank	e^*	$\mathbf{a}^*(0.98)$	rank	e^*	$\mathbf{a}^*(0.97)$	rank	e^*	$\mathbf{a}^*(0.96)$	rank	e^*
Japan	0.948	11		0.952	13		0.953	14		0.994	14		0.989	14	
Kenya	-1.448	38		-1.462	43		-1.469	50		-1.456	51		-1.477	53	
Korea	-0.212	27		-0.220	31		-0.213	34		-0.172	35		-0.200	37	
Luxembourg	1.253	6		1.264	6		1.274	6		1.311	6		1.306	5	
Madagascar	-1.357	37		-1.369	42		-1.396	49		-1.356	50		-1.366	52	
Malawi	-1.950	41		-1.947	46		-1.963	53		-1.936	55		-1.971	57	
Mali			0.975			0.975			0.975	-1.920	54		-1.951	56	
Morocco	-0.761	31		-0.783	35		-0.774	41		-0.724	42		-0.725	44	
Netherlands			0.99	1.192	8		1.197	8		1.247	8		1.239	8	
Nigeria			0.995	-1.271	41		-1.283	48		-1.253	49		-1.279	51	
Norway	1.090	9		1.087	10		1.093	11		1.135	12		1.138	11	
Pakistan			0.99			0.98	-0.727	40		-0.698	41		-0.699	43	
Panama	-0.076	24		-0.071	28		-0.074	31		-0.048	31		-0.057	32	
Paraguay			0.985			0.975			0.975	-0.110	33		-0.115	34	
Peru	-0.084	25		-0.101	29		-0.089	32		-0.059	32		-0.076	33	
Philippines			0.99			0.985	-0.343	36		-0.306	37		-0.314	39	
Poland			0.995			0.985	0.408	25		0.445	24		0.468	24	
Portugal	0.465	17		0.448	20		0.467	22		0.501	22		0.519	23	
Senegal	-1.160	36		-1.162	40		-1.170	47		-1.141	48		-1.165	50	
Spain			0.99			0.985	0.849	16		0.897	16		0.904	16	
Sri Lanka	-0.653	30		-0.644	34		-0.637	39		-0.602	40		-0.613	42	
U.R. Tanzania			0.995	-2.010	47		-2.019	54		-1.996	56		-2.012	58	
Tunisia			0.995			0.985	-0.296	35		-0.260	36		-0.264	38	
U.K.			0.99			0.985	1.107	10		1.147	10		1.148	10	
U.S.	1.456	1		1.465	1		1.466	1		1.507	1		1.507	1	
Uruguay	0.540	14		0.547	17		0.550	19		0.579	19		0.575	20	
Venezuela	0.527	15		0.536	18		0.536	20		0.540	20		0.539	21	
Yugoslavia			0.995	0.349	23		0.356	26		0.395	26		0.395	27	
Zambia	-1.810	40		-1.794	45		-1.807	52		-1.745	53		-1.768	55	
Zimbabwe			0.995			0.985	-1.131	45		-1.110	46		-1.136	48	

A test of HARP(0.99) was conducted next and 48 countries were found to share common approximate homothetic preferences.[7] The Ideal Afriat Index and implied country rankings for $e^* = 0.99$ are shown in columns 5 and 6. This process was continued until all countries except the Dominican Republic were found to share common approximate homothetic preferences. The Afriat efficiency index required to show the 59 countries being rationalised by an approximate homothetic utility function was $e^* = 0.96$. This implies a level of consumer optimisation error of 4.2 percent. To bring the Dominican Republic into the set of homothetic countries would require $e^* = 0.94$; this implies a level of optimisation error which is greater than the 5 percent which has been suggested as an appropriate upper limit.

One of the problems with the use of the standard Ideal Afriat Index is that it is not possible to assess the significance of pairwise rankings between particular countries. For example, in column 2 of Table 4.3 the Ideal Afriat Index for Denmark is 1.275 and the welfare level for Germany is 1.273. The question is: how confident are we that Denmark is truly ranked higher than Germany? While the use of the Approximate Ideal Afriat Index does not enable the construction of confidence intervals around the welfare indexes, it can be used to identify pairwise rankings which change under differing assumptions regarding the level of consumer optimisation error.

In Figure 4.11 the rankings of the 42 countries which satisfied HARP(1) are shown under differing assumptions regarding the level of e^*. Note that the countries are arranged along the horizontal axis according to their ranking implied by the standard Ideal Afriat Index (e.g., that shown in column 3 in Table 4.3). This figure can be used to easily identify those countries which experience a change in their pairwise rankings under differing levels of the Afriat efficiency index (a "dip" in a line in the figure is evidence of such a change). For example, while Denmark is ranked higher than Germany with $e^* = 1$, this ranking is reversed with an Afriat efficiency index as high as 0.99 (which implies only 1 percent optimisation error). One can therefore conclude that Denmark and Germany should perhaps be ranked as equivalent. Of the 42 countries found to satisfy HARP(1), the pairwise ranking between four other countries (Belgium and Luxembourg, and El Salvador and Bolivia) change under differing assumptions regarding the level of e^*.

[7] Note that the iterative procedure used to select the 48 countries started with an initial subset of the 42 countries which satisfy HARP(1), not the subset of four countries used previously.

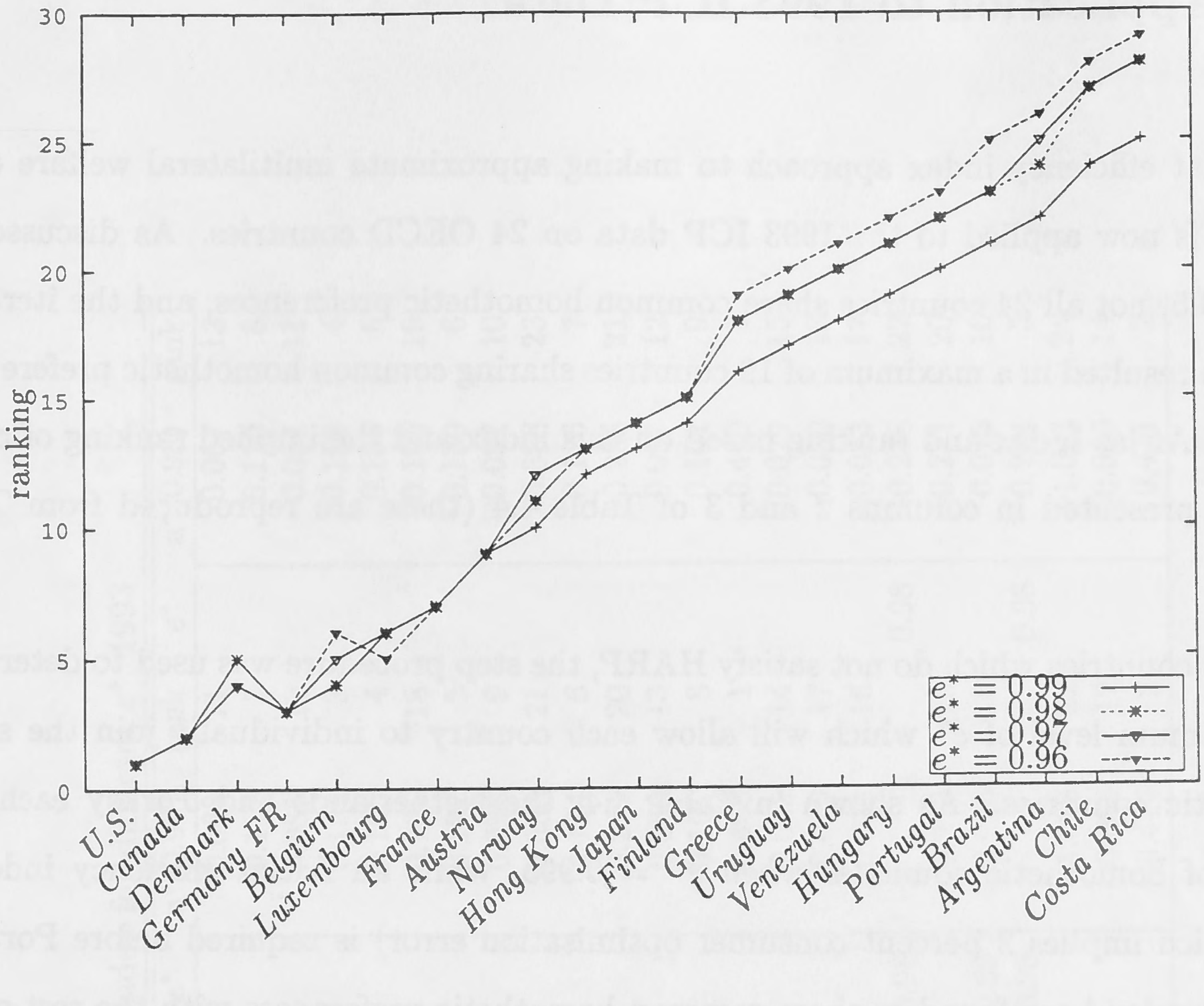

Figure 4.11: True welfare rankings calculated for different e^*, 1980

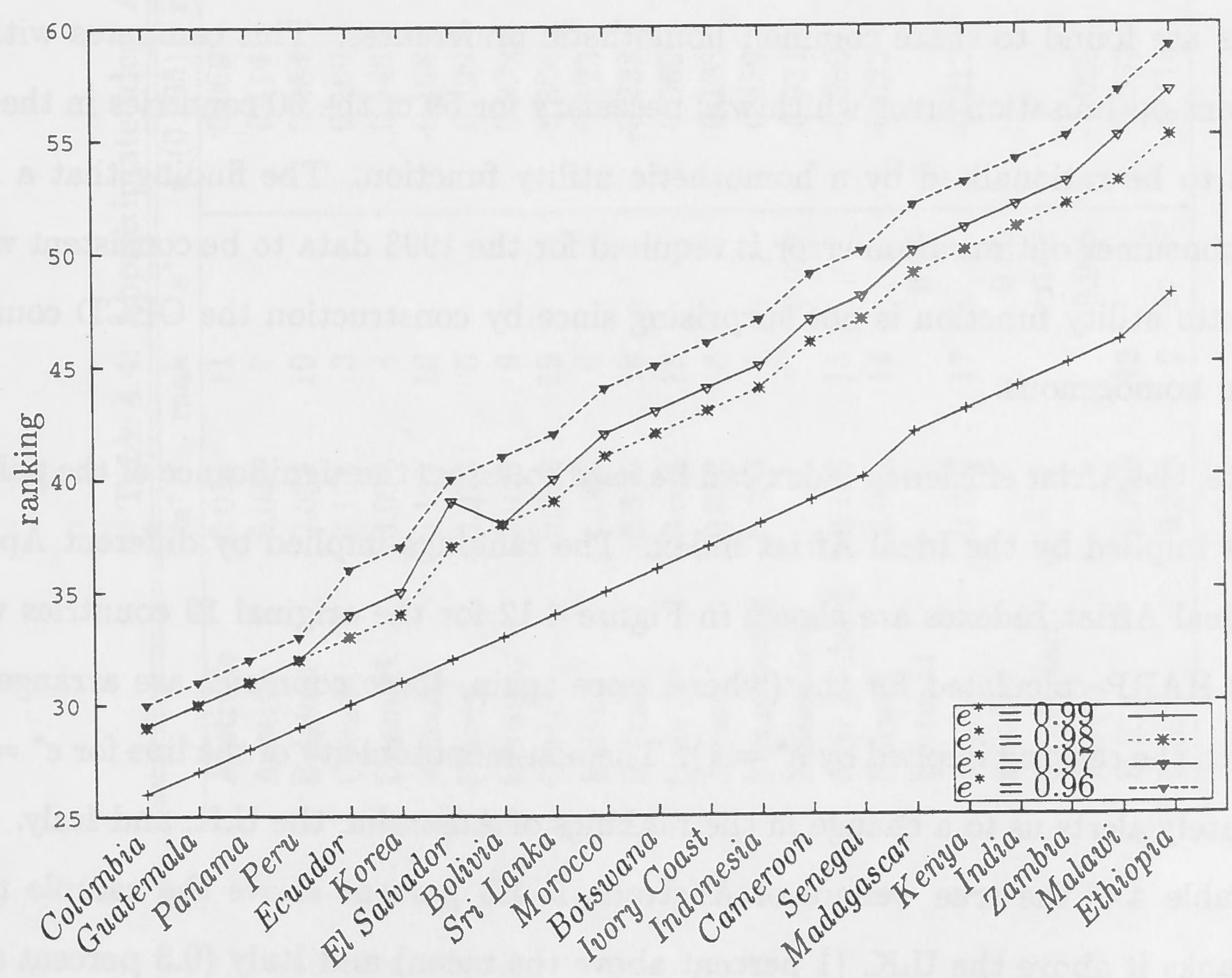

Figure 4.11: True welfare rankings calculated for different e^*, 1980 (cont.)

4.5 Application to 1993 ICP Data

The Afriat efficiency index approach to making approximate multilateral welfare comparisons is now applied to the 1993 ICP data on 24 OECD countries. As discussed in Section 3.5, not all 24 countries share common homothetic preferences, and the iterative approach resulted in a maximum of 19 countries sharing common homothetic preferences; the Ideal Afriat Index and ranking based on this index and the implied ranking of countries are presented in columns 2 and 3 of Table 4.4 (these are reproduced from Table 3.3).

For the 5 countries which do not satisfy HARP, the step procedure was used to determine the maximum level of e^* which will allow each country to individually join the set of homothetic countries. As shown in Table 4.4, the Netherlands and Turkey each join the set of homothetic countries when $e^* = 0.995$, while an Afriat efficiency index of 0.98 (which implies 2 percent consumer optimisation error) is required before Portugal and Switzerland are found to share common homothetic preferences with the rest of the OECD countries. Thus, with only 2 percent consumer optimisation error, all 24 OECD countries are found to share common homothetic preferences. This compares with the 4.2 percent optimisation error which was necessary for 59 of the 60 countries in the 1980 data set to be rationalised by a homothetic utility function. The finding that a lower level of consumer optimisation error is required for the 1993 data to be consistent with a homothetic utility function is not surprising since by construction the OECD countries are more homogenous.

As before, the Afriat efficiency index can be used to assess the significance of the pairwise rankings implied by the Ideal Afriat index. The rankings implied by different Approximate Ideal Afriat Indexes are shown in Figure 4.12 for the original 19 countries which satisfied HARP calculated for the (where, once again, these countries are arranged according to the ranking implied by $e^* = 1$). The non-monotonicity of the line for $e^* = 0.98$ immediately alerts us to a change in the rankings of Australia, the U.K. and Italy. From $\mathbf{a}^*$ in Table 4.4, the true welfare of Australia is 1.3 percent above the sample mean, which ranks it above the U.K. (1 percent above the mean) and Italy (0.3 percent above the mean). However, with $e^* = 0.98$ the welfare levels of these three countries are almost identical at approximately 7 percent above the mean, and when $\mathbf{a}^*(0.98)$ is reported to

Table 4.4: Approximate Ideal Afriat Index for different e^*, 1993

	$\mathbf{a}^*$	rank	e^*	$\mathbf{a}^*(0.995)$	rank	e^*	$\mathbf{a}^*(0.99)$	rank	e^*	$\mathbf{a}^*(0.98)$	rank
Australia	0.013	11		0.069	11		0.068	11		0.067	13
Austria	0.058	7		0.114	7		0.112	7		0.115	8
Belgium	0.028	10		0.084	10		0.083	10		0.084	11
Canada	0.105	3		0.160	3		0.159	3		0.158	4
Denmark	0.094	4		0.145	4		0.146	4		0.135	5
Finland	-0.160	16		-0.108	17		-0.105	18		-0.105	19
France	0.073	5		0.129	5		0.127	5		0.130	6
Germany	0.043	9		0.099	9		0.098	9		0.097	10
Greece	-0.362	19		-0.304	20		-0.306	21		-0.308	23
Iceland	0.067	6		0.123	6		0.124	6		0.126	7
Ireland	-0.300	18		-0.238	19		-0.241	20		-0.235	21
Italy	0.003	13		0.059	13		0.060	13		0.067	12
Japan	0.052	8		0.108	8		0.105	8		0.101	9
Luxembourg	0.400	1		0.459	1		0.458	1		0.440	1
Netherlands			0.995	0.031	14		0.029	14		0.029	15
New Zealand	-0.133	15		-0.080	16		-0.079	17		-0.080	18
Norway	-0.084	14		-0.023	15		-0.023	16		-0.013	17
Portugal			0.98			0.98			0.98	-0.238	22
Spain	-0.267	17		-0.212	18		-0.213	19		-0.211	20
Sweden			0.99			0.99	0.016	15		0.015	16
Switzerland			0.98			0.98			0.98	0.244	3
Turkey			0.995	-1.098	21		-1.098	22		-1.093	24
U.K.	0.010	12		0.067	12		0.065	12		0.066	14
U.S.	0.359	2		0.415	2		0.414	2		0.411	2

4 decimal places, Italy is in fact ranked above Australia and the U.K. The finding that with only 2 percent consumer optimisation error that the pairwise rankings between Australia, Italy and the U.K. change is evidence that these three countries should perhaps be ranked as equivalent.

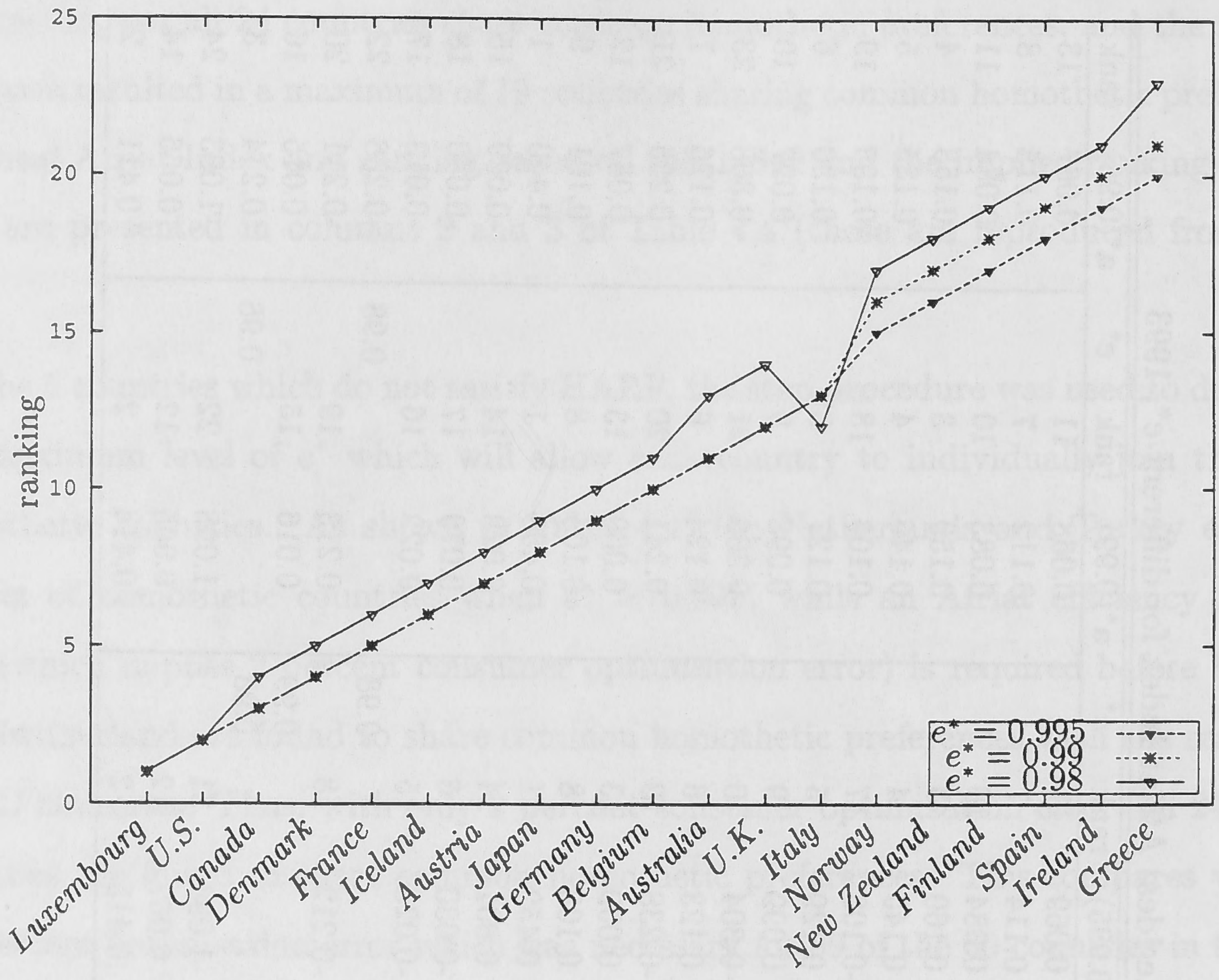

Figure 4.12: True welfare rankings calculated for different e^*, 1993

4.6 Conclusions

One of the criticisms of the RP approach to welfare measurement is that it is an "all or nothing" approach: the data either satisfy the RP test exactly or else they don't and the hypothesis of utility maximisation is rejected. In the context of the existence and construction of multilateral true welfare indexes, the HARP test of common homothetic preferences does not allow for the fact that consumption data are likely to be measured with error, and also consumers are unlikely to optimise without error. In certain contexts, it may be sufficient to simply find the largest set of observations which satisfy HARP and construct multilateral true indexes for that set. However, in a cross-country context, it is undesirable to simple drop countries which do not satisfy homotheticity from the analysis, and it is thus important to investigate methods for making utility consistent comparisons for all countries in a given data set.

In this chapter, two methods for making approximate multilateral true welfare comparisons were reviewed. The first was the Afriat envelope approach, which can be used to impute utility bounds for those countries which do not share common homothetic preferences. While the Augmented Ideal Afriat Index can be used to compare the welfare of countries in a data set that does not satisfy HARP, a problem with this approach is that the source of the failure of common preferences is not identified. In effect, the data are taken as if they were measured without error and there is no consumer optimisation error.

The Afriat efficiency index approach, in contrast, can be used to specifically incorporate different levels of consumer optimisation error into welfare comparisons. The Afriat efficiency index approach was applied to 1980 ICP data on 60 countries and it was found that while 42 countries share common homothetic preferences using the standard test of HARP, 59 countries satisfy HARP when 4.2 percent consumer optimisation error is allowed for. The Approximate Ideal Afriat Index was constructed and it was found that with only moderate levels of consumer optimisation error, the rankings of three pairs of countries change. It was suggested that these countries should perhaps be ranked equivalent, since their pairwise rankings change under differing (moderate) levels of e^*. All 24 OECD countries in the 1993 data set satisfy HARP when 2 percent optimisation error is allowed (while 19 satisfy HARP with zero optimisation error). Further, the pairwise rankings of Italy, Australia and the U.K. were found to indeterminate in that they change with different levels of e^*.

The Afriat efficiency index method for constructing approximate multilateral true welfare indexes can be used in other contexts. In Chapter 5 the method is used in the construction of an approximate multilateral true marginal welfare index, and in Chapter 6 the method is applied in the context of leisure-inclusive welfare comparisons.

Chapter 5

Multilateral True Marginal Indexes

5.1 Introduction

> "Empirical experience is abundant that the Santa Claus hypothesis of homotheticity in tastes and in technical change is quite unrealistic. Therefore, we must not be bemused by the undoubted elegances and richness of the homothetic theory...we must accept the sad facts of life, and be grateful for the more complicated procedures economic theory devises." P.A. Samuelson and S. Swamy (1974, p.592).

In making cross-country comparisons of welfare, it is intuitively desirable that welfare measurements not depend on which country is being used as the base in the comparison. As shown in Chapter 1, a welfare index that is invariant to the price vector used in its construction only exists when preferences are homothetic and for this reason, homotheticity has been integral in the construction of utility-consistent international comparisons of welfare. In Chapter 3, HARP was used to test 1980 and 1993 ICP data for the existence of a multilateral true welfare index, and for those countries sharing common homothetic preferences, welfare comparisons were conducted using the Ideal Afriat Index of Dowrick and Quiggin (1997).

However, homotheticity is a very restrictive assumption and may not be satisfied by all countries being studied. Further, homotheticity implies that in consumption space, the (linear) expansion paths originate from a single point, the origin, and consequently for any given relative price vector budget shares are constant across income levels (income elasticities are constrained to one). Individuals in rich and poor countries are therefore restricted to devote the same share of their budget to food, for example, and this contradicts empirical evidence that food is a necessity and its share in the budget decreases as income increases. Authors such as Varian (1983) and Manser and McDonald (1988) have found that U.S. time series consumption data satisfy HARP, however evidence from econometric studies (such as that by Blackorby, Boyce, and Russell (1978)) has tended to reject homotheticity for aggregate data.[1]

It therefore appears that an investigation of the construction of true indexes in the absence of homotheticity is warranted, especially in the context of cross-country comparisons of welfare. In the present chapter, a multilateral true marginal index is defined and constructed using ICP data. The multilateral true marginal index proposed is a direct analog of the Ideal Afriat Index, except it measures the utility gained from *marginal consumption* (consumption in excess of a predetermined minimum subsistence bundle, denoted γ), and for this reason it is named the Ideal Afriat Marginal Index, conditional on γ. The Ideal Afriat Marginal Index is constructed under the less restrictive assumption of affine-homothetic preferences, that is, the linear expansion paths originate from a particular point (the subsistence bundle), not necessarily the origin.

Affine homotheticity is still restrictive in terms of implied economic behaviour, however it is a significant relaxation from homotheticity. Under affine homotheticity the income elasticities for *marginal consumption* are constrained to one, but the income elasticities for *consumption* may differ from one (and thus goods can be found to be necessities or luxuries). Homotheticity is in fact a special case of affine homotheticity (when $\gamma = \mathbf{0}$) and hence the Ideal Afriat Index is a special case of the Ideal Afriat Marginal Index.[2] Moreover, it is shown that affine homotheticity is the only generalisation from homotheticity which is consistent with one being able to make multilaterally consistent (marginal) welfare comparisons.

In Chapter 1, it was argued that without complete knowledge of preferences, the Allen welfare index is unobservable and this leads to indeterminateness in that there is a range within which the index can lie (and any index within this range itself qualifies as a true

[1] One may question, however, an approach which uses a demand system derived from a particular preference structure since it really involves a test of the joint hypothesis that the functional form is correct and the data do not satisfy homotheticity.

[2] Further, since homotheticity is a special case of affine homotheticity, it is to be expected that in any data set more observations will be shown to be consistent with affine homotheticity.

index). The use of revealed preference methods can tighten the bounds to the true index, however the unique welfare index can never be exactly determined and there will always exist a family of homothetic functions that are consistent with any given data set that satisfies HARP. One of the main contributions of this chapter is to identify that there are really *two* sources of indeterminateness in the construction of utility consistent welfare comparisons. For any given data set satisfying affine homotheticity, there will be a set denoted $\mathcal{G}$ containing an infinite number of γs which, by definition, are consistent with common affine homothetic preferences, and there will be a family of affine homothetic functions consistent with each γ in $\mathcal{G}$.[3]

The approach in this chapter, therefore, is not to attempt to estimate a single minimum subsistence consumption bundle for use in the construction of the Ideal Afriat Marginal Index. The reason for this is that there is no empirical reason to base welfare comparisons on just one of the bundles in $\mathcal{G}$. Rather, for a given data set that satisfies HARP, an iterative approach is used to find a sample of γs and the *bounds* to the Ideal Afriat Marginal Index are then calculated. These bounds to the true marginal welfare index are then compared to the Ideal Afriat Index.

The Ideal Afriat Index and Ideal Afriat Marginal Indexes both measure the true (marginal) welfare of each country relative to the sample mean. It is conjectured that as soon as we "move off the origin" and begin using the Ideal Afriat Marginal Index in welfare comparisons, for "richer" countries, this index will be higher than Ideal Afriat Index, and the reverse will be true for "poorer" countries. We are therefore able to use the Ideal Afriat Index and the bounds to the Ideal Afriat Marginal Indexes to identify poor and rich countries. In particular, a country is defined as rich if its true welfare as measured by the Ideal Afriat Index is equal to the lower bound to its true marginal welfare as measured by the Ideal Afriat Marginal Indexes. Similarly, a country is defined as poor if the Ideal Afriat Index is equal to the upper bound of the Ideal Afriat Marginal Indexes.

Results in Chapter 4 on the construction of approximate welfare comparisons are extended to the construction of the Approximate Ideal Afriat Marginal Index and applied to the ICP data. For the 1980 data, it is found that five of the pairwise country rankings implied by the Approximate Ideal Afriat Index are contradicted by the bounds to the

[3] If the data also satisfy HARP, then $\gamma = \mathbf{0}$ will be one of these bundles.

Approximate Ideal Afriat Marginal Indexes. That is, while the use of $\gamma = \mathbf{0}$ results in particular country rankings, there exist other minimum subsistence consumption bundles in $\mathcal{G}$ which result in these rankings being reversed. The conclusion is that these 5 pairs of countries should perhaps be considered equivalent. It is also found that of the 59 countries for which (approximate) welfare comparisons can be made, 17 (or 28.8 percent) are poor. With the 1993 ICP data on 24 OECD countries, several inconsistencies in the country rankings implied by the Ideal Afriat Index are also identified. It is further found that 5 countries (or 20.8 percent of the sample) are poor, although it should be emphasised that these countries are poor relative to the other OECD countries.

The structure of the chapter is as follows. In Section 5.2, previous work on constructing marginal indexes is reviewed. Results on the construction of multilateral true marginal welfare indexes are presented in Section 5.3, and a method for constructing bounds to the Ideal Afriat Marginal Indexes is proposed. In Section 5.4, the method is applied to a 5 country example data set, and results on classifying countries as either rich or poor and estimating income elasticities of demand are presented. Sections 5.5 and 5.6 present applications of the method to 1980 and 1993 ICP data, respectively. Section 5.7 concludes this chapter.

5.2 Previous Work on Constructing Marginal Indexes

As is the case with most topics associated with index number theory, the construction of marginal index numbers has been thoroughly investigated in the literature. Wald's (1939) "New Formula" for the cost of living index is based on an assumption that expansion paths are linear (but not necessarily from the origin). With two observed consumption bundles for each relative price vector, the expansion path is fully determined; Wald (1939) used this result to derive a formula to measure the cost of living between two periods which is consistent with linear expansion paths (and since a general quadratic utility function generates linear expansion paths which do not pass through the origin, the index is consistent with the existence of such a utility function).

Sidney N. Afriat showed in various articles that Wald's "New Formula" could be derived using a generalisation of the result of Byushgens (1925) that the Fisher index is exact for homogeneous quadratic preferences.[4] Afriat raised a major concern with the use of the Wald's "New" index and also the Fisher index. Afriat was not concerned that in both cases preferences were assumed to be of a of a particular form (in the case of the Fisher index, homogeneous quadratic, and in the case of Wald's "New" formula, general quadratic), rather that there was no explicit *testing* of whether such preferences held for the observations for which the welfare or price comparison was being made.[5]

In developing his critique Afriat employed quasi homotheticity (which is characterised by the Gorman Polar Form expenditure function). A feature of quasi-homothetic preferences is linear expansion paths and hence a general quadratic utility function is an example of quasi homotheticity (as is homogeneous quadratic utility, to the extent that it implies homothetic preferences, and homotheticity is a special case of quasi homotheticity). Afriat used quasi homotheticity to define a money metric true index which he called the marginal index and he further showed that Wald's index was in fact a marginal index and how if the linear expansion paths originated from the origin, then the marginal index was identical to the Fisher index.[6]

Afriat also gave several important new results on the construction of marginal indexes.

[4] Afriat (1987) contains a summary of this research.

[5] As a multilateral generalisation of the Fisher index, this same critique can be leveled at the EKS index, or indeed any superlative index.

[6] Most of Afriat's work on marginal indexes related to the marginal price index, but the results are easily adapted to the construction of a marginal welfare index.

First, he showed that given predetermined linear expansion paths (for example via knowledge of two consumption bundles for every set of relative prices), revealed preference could be used to test whether the data were consistent with the preferences underlying the linear expansion paths. Second, Afriat constructed bounds to the bilateral true marginal indexes (and showed how in the special case of homotheticity with $N = 2$ these bounds collapse to the Laspeyres and Paasche indexes).

Finally, Afriat showed that in the (more realistic) situation of knowledge of only one consumption point per expansion path, it was still possible to test for the existence of a true marginal index and construct bounds to this true index. His "Four point theorem" was therefore a completely nonparametric method for constructing bounds to welfare indexes - all that was required was an assumption of general quadratic preferences and the relevant consumption data. In this sense, Afriat's "Four point theorem" is a more general version of his own results on using nonparametric methods to constructing homothetic welfare indexes, later developed by Varian (1983) and used in a multilateral context by Dowrick and Quiggin (1997). However, Afriat showed that in general, his nonparametric approach only worked for four countries (hence the name of the formula) and since multilateral comparisons of welfare are usually conducted for more than four countries, the "Four point theorem" is of limited use in the present context.

Afriat's work on marginal indexes provides a fertile ground of ideas which apparently have not been extensively researched. Afriat (1977) calculated bilateral marginal price indexes using 1969 Canadian data, but while Dowrick and Quiggin (1997) have extended Afriat's early work on true indexes (which by definition require homotheticity) to multilateral comparisons, there does not appear to be any work on investigating the use of marginal indexes in the multilateral context.[7]

[7]Afriat (1972, p.63) notes that his results on marginal indexes may be extended to more than two countries.

5.3 Affine Homotheticity and the Multilateral True Marginal Welfare Index

It was shown in Chapter 1 that homotheticity is necessary and sufficient for the existence of a unique welfare index. In Chapter 3, several methods for testing the consistency of a given finite set of data with homotheticity were presented; of these tests, the test of HARP is the most easily implemented. A particular multilateral true index, the Ideal Afriat Index of Dowrick and Quiggin (1997) has been suggested as appropriate index for making cross-country welfare comparisons and in Chapter 3, this index was constructed for the 1980 and 1993 ICP data.

In this section, the above results on the existence and construction of multilateral true indexes are generalised. A multilateral true marginal welfare index which measures the utility gained from *marginal* or *supernumerary consumption* (consumption in excess of a predetermined minimum subsistence bundle) is defined and a test for its existence is presented. It is shown that affine homotheticity is necessary and sufficient for the existence of a marginal welfare index, and that an easily implemented test of affine homotheticity involves the Affine Homothetic Axiom of Revealed Preference (AHARP). A particular multilateral true marginal welfare index, the Ideal Afriat Marginal Index is proposed and later in the chapter, the bounds to this index are constructed using the ICP data.

5.3.1 Quasi-homothetic preferences

Afriat (1972, 1977, 1987) investigated the construction of marginal indexes which are defined under quasi-homothetic preferences. Quasi homotheticity is characterised by the Gorman Polar Form (GPF) expenditure function:

$$e(U, \mathbf{p}) = a(\mathbf{p}) + b(\mathbf{p})U, \tag{5.1}$$

where $a(\mathbf{p})$ and $b(\mathbf{p})$ are positive and homogeneous of degree 1 functions of prices. The GPF has the economic interpretation that $a(\mathbf{p})$ is the fixed cost of attaining the base utility level (normalised to be zero) and thus is the cost of purchasing a subsistence bundle of goods, and $b(\mathbf{p})$ is marginal price of attaining utility U above the subsistence level.

Applying Shephard's lemma to (5.1) gives the Hicksian demand function for good l:

$$h_l(U, \mathbf{p}) = \nabla_l a(\mathbf{p}) + \nabla_l b(\mathbf{p})U, \tag{5.2}$$

where $\nabla_l a(\mathbf{p}) = \partial a(\mathbf{p})/\partial p_l$ and $\nabla_l b(\mathbf{p}) = \partial b(\mathbf{p})/\partial p_l$. Using (5.1) to substitute for U in the above gives the Marshallian demand function for good l:

$$q_l(x, \mathbf{p}) = \nabla_l a(\mathbf{p}) + \nabla_l b(\mathbf{p})/b(\mathbf{p})[x - a(\mathbf{p})].$$

Thus, quasi homotheticity implies linear expansion paths and Engel curves. A demand system is theoretically plausible vis-a-vis aggregate data only if they can be generated from an aggregate preference ordering (i.e. are consistent with the existence of a utility maximising representative consumer). It can be shown that linear Engel curves (for which quasi homotheticity is both necessary and sufficient) are the minimum requirement for exact aggregation consistency (see, for example, Deaton and Muellbauer (1980)). For this reason, demand systems derived from the GPF (such as the Linear Expenditure System, LES) have been used extensively in empirical research using aggregate data.[8]

The expansion paths for quasi-homothetic preferences are illustrated for the two-good case in Figure 5.1. $EP(\mathbf{p}^1)$ and $EP(\mathbf{p}^2)$ are expansion paths at prices $\mathbf{p}^1$ and $\mathbf{p}^2$, respectively, and the subsistence level consumption bundles are $\nabla a(\mathbf{p}^1)$ and $\nabla a(\mathbf{p}^2)$ (found by setting $U = 0$ in (5.2)).[9] Note that since $\nabla a(\mathbf{p}^1)$ is not in the positive orthant in this particular example, $EP(\mathbf{p}^1)$ is piecewise linear from the origin to the point $\overline{\mathbf{q}}$.

Homotheticity is a special case of the GPF, generated by $a(\mathbf{p}) = 0$. With homotheticity the base utility curve degenerates to a single point - the origin - and all expansion paths are rays (Figure 5.2). An intermediate special case is *affine homotheticity* which is generated by:

$$a(\mathbf{p}) = \sum_l \gamma_l p_l,$$

in which case the base utility curve degenerates to the point $\gamma = (\gamma_1, \ldots, \gamma_K)$, not necessarily the origin (Figure 5.3).

---

[8] For an example of estimating the LES using aggregate cross-country data, see Lluch, Powell, and Williams (1977).

[9] In general, $\gamma^r = \nabla a(\mathbf{p}^r)$.

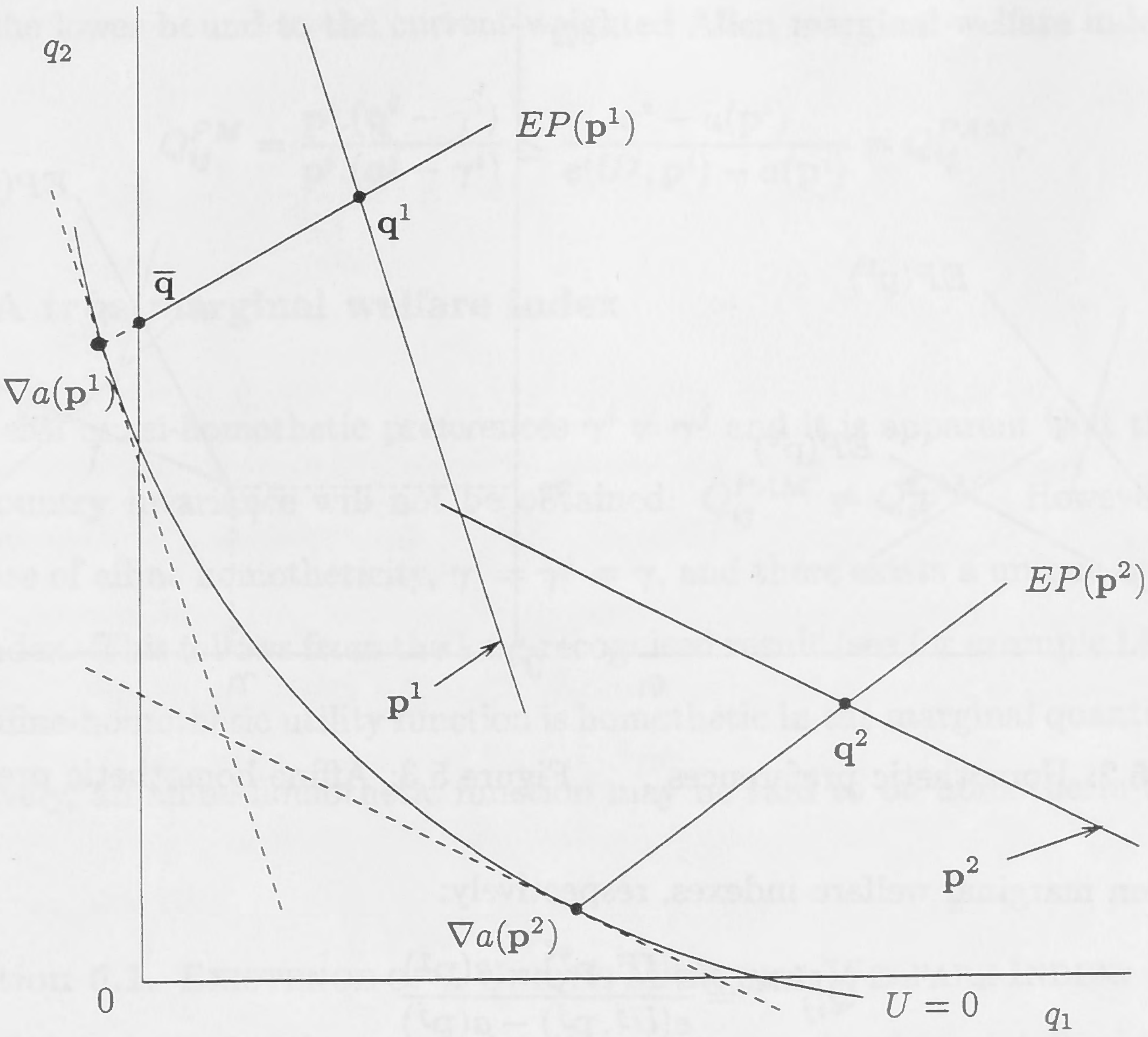

Figure 5.1: Quasi-homothetic preferences

5.3.2 Marginal welfare indexes

Analogous to the discussion in Section 1.4.2, there are three marginal welfare indexes which can be constructed. However, the Allen, Konüs and Malmquist marginal welfare indexes coincide when preferences are affine homothetic, and hence the following discussion is limited to the Allen marginal index.

First assume preferences are quasi homothetic. The Allen marginal welfare index at reference prices $\mathbf{p}^r$ is defined (using 5.1):

$$Q_{ij}^{AM,r} = \frac{e(U^i, \mathbf{p}^r) - a(\mathbf{p}^r)}{e(U^j, \mathbf{p}^r) - a(\mathbf{p}^r)}.$$

The Allen marginal welfare index gives the fraction of the cost of attaining country j's marginal utility level required to attain country i's marginal utility level at the reference prices. The natural candidates for $\mathbf{p}^r$ are $\mathbf{p}^j$ and $\mathbf{p}^i$, giving the Laspeyres-Allen and

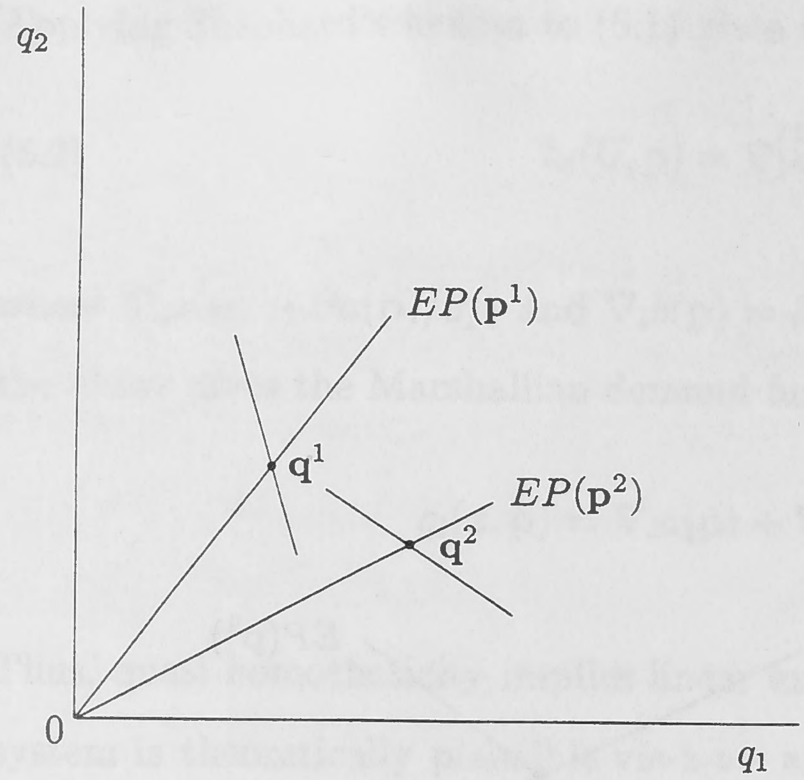

Figure 5.2: Homothetic preferences

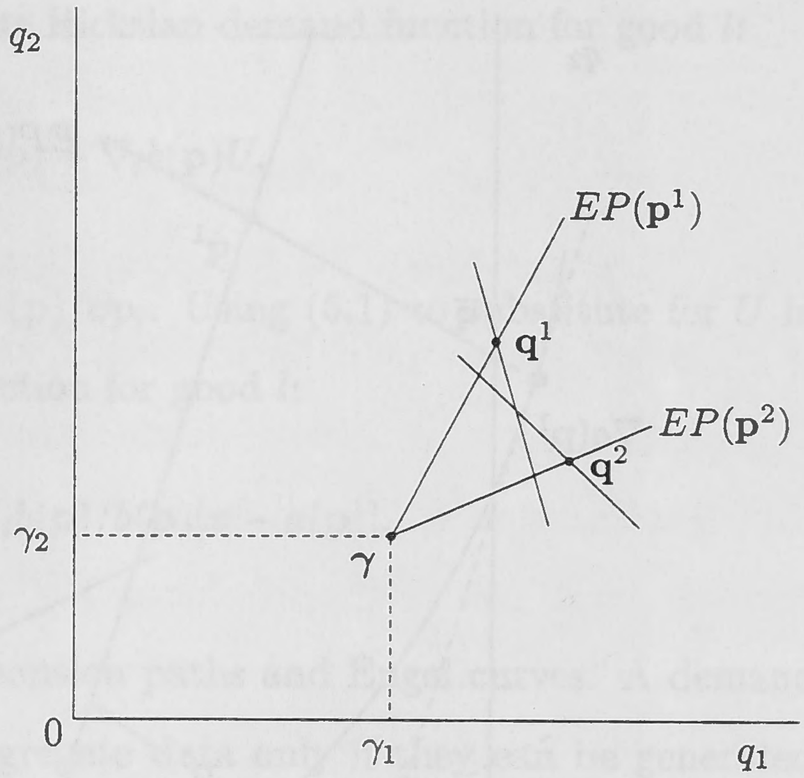

Figure 5.3: Affine-homothetic preferences

Paasche-Allen marginal welfare indexes, respectively:

$$Q_{ij}^{LAM} = \frac{e(U^i, \mathbf{p}^j) - a(\mathbf{p}^j)}{e(U^j, \mathbf{p}^j) - a(\mathbf{p}^j)}$$
$$Q_{ij}^{PAM} = \frac{e(U^i, \mathbf{p}^i) - a(\mathbf{p}^i)}{e(U^j, \mathbf{p}^i) - a(\mathbf{p}^i)}.$$

The Allen marginal welfare indexes Q_{ij}^{LAM} and Q_{ij}^{PAM} are not directly observable, however, analogous to the discussion in Section 1.4.3, we can construct classical and fixed-weight bounds to these indexes.[10] The Laspeyres and Paasche marginal quantity indexes are defined, respectively:

$$Q_{ij}^{LM} = \frac{\mathbf{p}^j.(\mathbf{q}^i - \gamma^j)}{\mathbf{p}^j.(\mathbf{q}^j - \gamma^j)}$$
$$Q_{ij}^{PM} = \frac{\mathbf{p}^i.(\mathbf{q}^i - \gamma^i)}{\mathbf{p}^i.(\mathbf{q}^j - \gamma^i)}.$$

The bundle $(\mathbf{q}^i - \gamma^j)$ is one way of achieving $U^i = u(\mathbf{q}^i - \gamma^j)$, but not necessarily the cheapest when prices are $\mathbf{p}^j$; hence $\mathbf{p}^j.(\mathbf{q}^i - \gamma^j) \geq e(U^i, \mathbf{p}^j) - a(\mathbf{p}^j)$. By definition, $\mathbf{p}^j.(\mathbf{q}^j - \gamma^j) = e(U^j, \mathbf{p}^j) - a(\mathbf{p}^j) = x^j - a(\mathbf{p}^j)$. Hence:

$$Q_{ij}^{LM} = \frac{\mathbf{p}^j.(\mathbf{q}^i - \gamma^j)}{\mathbf{p}^j.(\mathbf{q}^j - \gamma^j)} \geq \frac{e(U^i, \mathbf{p}^j) - a(\mathbf{p}^j)}{x^j - a(\mathbf{p}^j)} = Q_{ij}^{LAM},$$

that is, the Laspeyres marginal quantity index is the upper bound to the base-weighted Allen marginal welfare index. Using similar reasoning the Paasche marginal quantity

[10] The classical bounds are redundant under affine-homothetic preferences, and hence are not discussed here.

index is the lower bound to the current-weighted Allen marginal welfare index:

$$Q_{ij}^{PM} = \frac{\mathbf{p}^i.(\mathbf{q}^i - \boldsymbol{\gamma}^i)}{\mathbf{p}^i.(\mathbf{q}^j - \boldsymbol{\gamma}^i)} \leq \frac{x^i - a(\mathbf{p}^i)}{e(U^j, \mathbf{p}^i) - a(\mathbf{p}^i)} = Q_{ij}^{PAM}.$$

5.3.3 A true marginal welfare index

With general quasi-homothetic preferences $\gamma^i \neq \gamma^j$ and it is apparent that the property of base-country invariance will not be obtained: $Q_{ij}^{PAM} \neq Q_{ij}^{LAM}$. However, with the special case of affine homotheticity, $\gamma^i = \gamma^j = \gamma$, and there exists a unique true marginal welfare index. This follows from the long-recognised result (see for example Lloyd (1979)) that an affine-homothetic utility function is homothetic in the marginal quantities $(\mathbf{q}-\boldsymbol{\gamma})$. Alternatively, an affine-homothetic function may be said to be homothetic to the point γ.

Proposition 5.1. Existence of a Unique Marginal Welfare Index*: There exists a unique index comparing the marginal welfare of country i and j and defined by:* $Q_{ij}^{LAM} = Q_{ij}^{PAM} = U^i/U^j$*, if and only if preferences are affine homothetic.*

Proof: Analogous to the proof to Proposition 1.2.

Having thus shown that there exists a unique marginal welfare index when preferences are affine-homothetic, the next step is to devise a test for the consistency of a given finite data set with affine homotheticity. In Proposition 3.1, it was shown that a necessary and sufficient condition for the consistency of a given finite set of data with homotheticity is the test of HARP. This leads to the following test for affine homotheticity:

Proposition 5.2. Common Affine Homothetic Preferences*: A finite set of demand data is consistent with the existence of common affine homothetic preferences if and only if there exists a quantity vector which can be defined as a minimum subsistence consumption bundle γ, that is:*

1. $\boldsymbol{\gamma} = \{\gamma_1, \gamma_2, \ldots, \gamma_K\} \geq \mathbf{0}$;

2. γ *is such that no country has negative marginal consumption of any good;*[11]

[11] While in theory it is possible to have negative marginal consumption, this is inconvenient since the test for common affine homothetic preferences involves taking logarithms of values.

3. *The data satisfy the Affine Homothetic Axiom of Revealed Preference (AHARP): For all distinct choices of indexes $(i, j, ..., m)$ we have:*

$$\{\frac{\mathbf{p}^i.(\mathbf{q}^j-\gamma)}{\mathbf{p}^i.(\mathbf{q}^i-\gamma)}\}\{\frac{\mathbf{p}^j.(\mathbf{q}^k-\gamma)}{\mathbf{p}^j.(\mathbf{q}^j-\gamma)}\}\ldots\{\frac{\mathbf{p}^m.(\mathbf{q}^i-\gamma)}{\mathbf{p}^m.(\mathbf{q}^m-\gamma)}\}\geq 1$$

PROOF: Analogous to the proof to Proposition 3.1.

Thus, a data set is consistent with affine homotheticity if there exists a quantity vector which can be defined as a minimum subsistence consumption bundle. From the above proposition, a quantity vector can be defined as a minimum subsistence bundle if all elements are non-negative, no country has negative marginal consumption and it is consistent with the data satisfying AHARP. For a given γ, AHARP can be easily tested by inputting a matrix of logarithms of Laspeyres marginal quantity indexes $\{L^m_{ij}\} = \log Q^{LM}_{ji}$ into Warshall's algorithm to construct a minimum path matrix defined over marginal quantities $\mathbf{M}^m$. If any of the diagonal elements of $\mathbf{M}^m$ are negative, then the data do not satisfy affine homotheticity.

With affine homotheticity, it is apparent that the unique marginal welfare index will be bounded by the fixed-weight marginal quantity indexes:

$$Q^{PM}_{ij} \leq \frac{U^i}{U^j} \leq Q^{LM}_{ij}.$$

This leads to the definition of the unique true marginal welfare index which can be used empirically.

Definition 5.1. UNIQUE TRUE MARGINAL WELFARE INDEX *For a given minimum subsistence consumption bundle γ, a unique true marginal welfare index is a set of numbers $A^m = (A^{m1}, A^{m2}, \ldots A^{mN})$ such that:*

$$Q^{PM}_{ij} \leq A^{mi}/A^{mj} \leq Q^{LM}_{ij} \quad \forall i, j = 1, \ldots, N.$$

5.3.4 The Ideal Afriat Marginal Index

The results on the existence and construction of multilateral true welfare indexes presented in Chapters 1 and 3 can be generalised to multilateral true marginal welfare indexes.

Definition 5.2. MULTILATERAL MARGINAL INDEX *A multilateral marginal index is a marginal index which satisfies the property of circularity.*

The Laspeyres and Paasche marginal quantity indexes do not satisfy circularity, and hence do not quality as multilateral marginal indexes.

Definition 5.3. MULTILATERAL TRUE MARGINAL INDEX *A multilateral true marginal index is a true marginal index which also has the property of being a multilateral marginal index.*

We are now in a position to define a particular multilateral true marginal welfare index which will be used in the empirical part of this chapter.

Definition 5.4. THE IDEAL AFRIAT MARGINAL INDEX: *For a given minimum subsistence consumption bundle γ, let $\mathbf{a}_m^+ \equiv (c^1, c^2, \ldots, c^N)$ and $\mathbf{a}_m^- \equiv (-r^1, -r^2, \ldots, -r^N)$ represent the vector of column means and the vector of negative row means, respectively, of the minimum path matrix defined over marginal quantities, $\mathbf{M}^m$. Define the Ideal Afriat Marginal Index as $\mathbf{a}_m^* \equiv (\mathbf{a}_m^+ + \mathbf{a}_m^-)/2$, the vector of overall means.*

The properties and interpretation of the Ideal Afriat Marginal Index are analogous to that of the Ideal Afriat Index (see Proposition 3.4). Finally, it can be shown that the Ideal Afriat Marginal Index is a generalisation of the Ideal Afriat Index.

Proposition 5.3. IDEAL AFRIAT MARGINAL INDEX AND IDEAL AFRIAT INDEX: *The Ideal Afriat Index is a special case of the Ideal Afriat Marginal Index.*

PROOF: This follows from the fact that homotheticity is a special case of affine homotheticity, that is, preferences are homothetic when $\gamma = \mathbf{0}$.

5.3.5 An approach for determining the bounds to γ

In Chapter 1, it was argued that without complete knowledge of preferences, the welfare index is unobservable and this leads to indeterminateness in that there is a range within which the welfare index can lie (and any index within this range itself qualifies as a true index). While the use of revealed preference methods can tighten the bounds to the true index, the unique welfare index can never be exactly determined because a whole family of homothetic functions will be consistent with any given data set that satisfies HARP. However, as shown above, homotheticity is a special case of affine homotheticity and the multilateral true welfare index is a special case of the multilateral true marginal welfare index where $\gamma = 0$. Hence, there are really *two* sources of indeterminateness in the construction of true utility consistent welfare comparisons. For any given data set satisfying AHARP, there will be an infinite number of γs (contained within a particular range) which, by definition, are consistent with common affine homothetic preferences[12] and there will be a family of affine homothetic functions consistent with each γ.

Definition 5.5. SET OF MINIMUM SUBSISTENCE BUNDLES *For a given finite set of demand data, the set of minimum subsistence bundles is denoted $\mathcal{G}$.*

For a set of data that satisfies AHARP, $\mathcal{G}$ will contain an infinite number of γs, contained within a particular range. For a set of data that is consistent with HARP, one of the vectors in $\mathcal{G}$ will be $\boldsymbol{\gamma} = \mathbf{0}$. Obviously, for a given data set there may not exist any γ; in that case $\mathcal{G}$ will be the empty set.

If we had complete knowledge of preferences, we would know the γ to be used in the construction of the multilateral true marginal welfare index. However, this is not the case, and for a given data set which satisfies AHARP, there is no empirical reason to choose a particular γ in $\mathcal{G}$ for use in the construction of the relevant Ideal Afriat Marginal Index. Hence, in using the Ideal Afriat Index for cross-country welfare comparisons we are in fact ignoring all the other elements of $\mathcal{G}$ which can be equally validly used to specify the family of affine-homothetic utility functions which rationalises the data. Rather than arbitrarily selecting a particular γ in $\mathcal{G}$ and constructing the appropriate Ideal Afriat Marginal Index, the approach proposed in this chapter is to empirically determine the bounds to $\mathcal{G}$, and thus construct bounds to the Ideal Afriat Marginal Indexes.

The bounds to $\mathcal{G}$ are determined in the following way. For a given data set of N countries that satisfies HARP, a quantity vector is randomly chosen from a uniform distribution, the boundaries of which are a vector of zeros and a vector of minimum quantities $\mathbf{q}^{min} = \min_{i \in N}[q_l^i : l = 1, \ldots K]$. This quantity vector is a "potential" γ, and hence it cannot result in any country having below subsistence consumption. The test of AHARP is then conducted using the selected quantity vector as a minimum subsistence consumption bundle, and if the test is satisfied then the quantity vector is placed in the set $\mathcal{G}$. This process is repeated until $\mathcal{G}$ contains enough elements so that we can be reasonably confident that γs close to the boundaries of $\mathcal{G}$ have been selected. Note that it essentially

[12] This assumes that the consumption quantities can be measured continuously.

an arbitrary decision as to when enough γs have been collected for the boundaries of $\mathcal{G}$ to be adequately known.[13]

Once enough elements of $\mathcal{G}$ have been collected using this process, the Ideal Afriat Marginal Index is constructed for each γ in $\mathcal{G}$. The range in which $\mathbf{a}_m^*$ and the implied rankings lie are then presented and compared with the Ideal Afriat Index (and implied rankings).[14]

---

[13]In a future version of this research, the process of finding the boundaries to $\mathcal{G}$ will be made more accurate via a grid search procedure.

[14]In a previous version of this chapter, the approach used was to select a particular γ in $\mathcal{G}$ for the construction of the Ideal Afriat Marginal Index. In particular, the mean γ was used. However, there are two problems with this approach. First, the selection of the mean γ for use in constructing true marginal welfare comparisons relies on the fact that $\mathcal{G}$ is a convex set (otherwise the mean γ may lie outside $\mathcal{G}$, that is, it wouldn't be a minimum subsistence consumption bundle). While it is this author's belief that $\mathcal{G}$ is a convex set, this has not been proved. Second, even if $\mathcal{G}$ is convex, as mentioned earlier, there is no reason to base the welfare comparisons on a single vector in $\mathcal{G}$.

5.4 Bounds to the Ideal Afriat Marginal Index - Example

The data used in this example consist of five observations on two goods - the same data that were used to construct the Ideal Afriat Index in Section 3.3.4. Using the approach outlined above, 100 elements of $\mathcal{G}$ were identified - these are drawn as dots in Figure 5.4 (note that since these data satisfy homotheticity, $\boldsymbol{\gamma} = \mathbf{0}$ is one of the elements of $\mathcal{G}$).

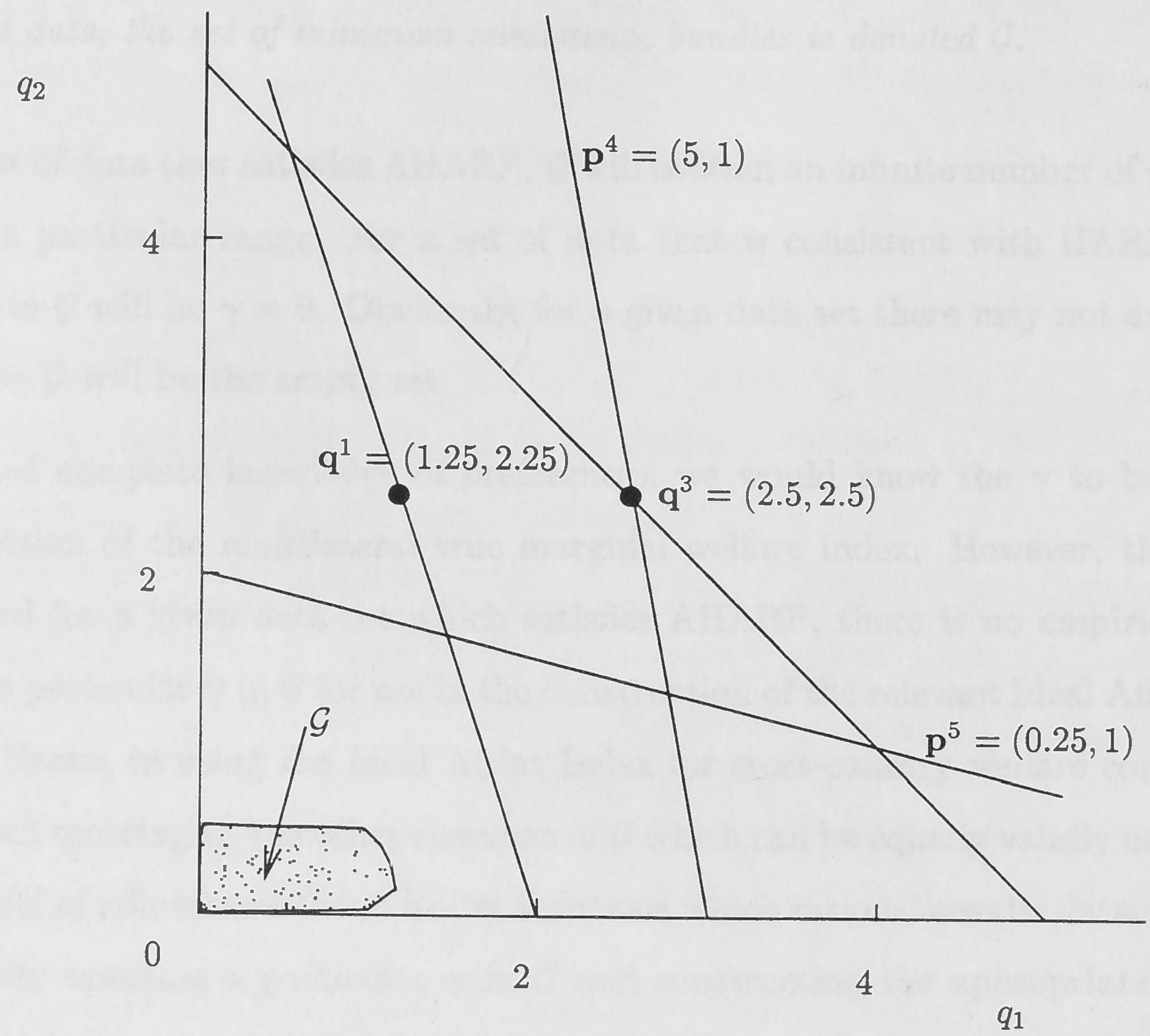

Figure 5.4: An estimate of $\mathcal{G}$ - five country homothetic demand data

Next, the Ideal Afriat Marginal Index was calculated for each $\boldsymbol{\gamma}$ in $\mathcal{G}$ - the ranges of these 100 $\mathbf{a}_m^*$s are presented in Table 5.1, as well as the Ideal Afriat Index (which is replicated from Table 3.1).

Table 5.1: Ideal Afriat Index and Ideal Afriat Marginal Index bounds, $N = 5$

	Country 1	Country 2	Country 3	Country 4	Country 5
$\mathbf{a}^*$	−0.479	0.510	−0.046	0.151	−0.136
min. $\mathbf{a}_m^*$	−0.819	0.510	−0.067	0.151	−0.192
max. $\mathbf{a}_m^*$	−0.479	0.656	−0.036	0.427	−0.136

5.4.1 A classification of poor and non-poor countries

A somewhat surprising feature of Table 5.1 is that for 4 of the 5 countries, the Ideal Afriat Index provides either the lower- or upper-bound to the range of the Ideal Afriat Marginal Index. In particular, for the countries ranked first and second according to $\mathbf{a}^*$ (countries 2 and 4), the Ideal Afriat Index is the lower bound to $\mathbf{a}^*_m$, while for two bottom ranked countries (countries 1 and 5), $\mathbf{a}^*$ provides an upper bound to $\mathbf{a}^*_m$. For the country ranked in the middle in terms of welfare (country 3), $\mathbf{a}^{*3}$ lies within the range of the Ideal Afriat Marginal Index.

On reflection, this is not such an unexpected result. The Ideal Afriat Index measures the true welfare of each country relative to the sample mean: as soon as we acknowledge the potential existence of a minimum subsistence consumption bundle (which is the same for all countries) and consequently begin using the Ideal Afriat Marginal Index (which is also calculated relative to the sample mean), it is conjectured that $\mathbf{a}^*_m$ will be higher than $\mathbf{a}^*$ for "richer" countries, and the reverse will be true for "poorer" countries. This observation can be formalised in the following definition.

Definition 5.6. RICH AND POOR COUNTRIES: *For a given data set that satisfies HARP, countries can be classified into three categories:*

1. *Country r is defined as rich if:* $\mathbf{a}^{*r} = \min_{g=1}^{G} a_m^{*r}[g]$;

2. *Country p is defined as poor if:* $\mathbf{a}^{*p} = \max_{g=1}^{G} a_m^{*p}[g]$;

3. *Country u is defined as neither rich nor poor if:* $\min_{g=1}^{G} a_m^{*u}[g] \leq \mathbf{a}^{*u} \leq \max_{g=1}^{G} a_m^{*u}[g]$,

where $a_m^{*i}[g]$ is the Ideal Afriat Marginal Index comparing the true marginal welfare of country i to the sample mean calculated using the gth minimum subsistence consumption bundle in $\mathcal{G}$, and the total number of vectors sampled from $\mathcal{G}$ is G.

Applying the above definition to the example data set, countries 2 and 4 are rich, country 3 is undetermined (neither rich nor poor), and countries 1 and 5 are poor. Below, this definition is used to classify the countries in the ICP data sets.

5.4.2 Calculating bounds to income elasticities of demand

With affine homotheticity, the income elasticity of *marginal consumption* will obviously be equal to one for all goods. However, the income elasticity of *consumption* for a particular commodity is not constrained to one and can be calculated as shown in the following proposition.

Proposition 5.4. INCOME ELASTICITY OF CONSUMPTION *For a given bundle in $\mathcal{G}$, denoted $\bar{\gamma}$, country i's income elasticity of consumption for good l is given by:*

$$\varepsilon_l^i = \frac{\bar{q}_l^i}{\bar{x}^i}\frac{x^i}{q_l^i},$$

where $\bar{q}_l^i = q_l^i - \bar{\gamma}_l$ and $\bar{x}^i = \mathbf{p}^i.\bar{\mathbf{q}}^i$.

PROOF: For a given $\bar{\gamma}$ in $\mathcal{G}$ we know that the income elasticity of marginal consumption of good l for country i:

$$\bar{\varepsilon}_l^i = \frac{\Delta\bar{q}_l^i}{\Delta\bar{x}^i}\frac{\bar{x}^i}{\bar{q}_l^i} = 1, \tag{5.3}$$

where $\bar{q}_l^i = q_l^i - \bar{\gamma}$, and "affine" income or expenditure is $\bar{x}^i = \mathbf{p}^i.\bar{\mathbf{q}}^i$. The objective is to calculate the income elasticity of consumption of good l:

$$\varepsilon_l^i = \frac{\Delta q_l^i}{\Delta x^i}\frac{x^i}{q_l^i}.$$

Assume that affine income has increased by a factor of ξ, that is $\Delta\bar{x}^i/\bar{x}^i = \xi - 1$. First, note that $\Delta q_l^i = \Delta\bar{q}_l^i$. Using (5.3), it is therefore apparent that:

$$\Delta q_l^i = \frac{\bar{q}_l^i}{\bar{x}^i}\Delta\bar{x}^i = \bar{q}_l^i(\xi - 1).$$

Similarly, since $\Delta x^i = \sum_l p_l^i \Delta q_l^i = \sum_l p_l^i \bar{q}_l^i(\xi - 1)$, the income elasticity of consumption for good l can therefore be calculated as:

$$\varepsilon_l^i = \frac{\bar{q}_l^i(\xi-1)}{\sum_l p_l^i\bar{q}_l^i(\xi-1)}\frac{x^i}{q_l^i} = \frac{\bar{q}_l^i}{\sum_l p_l^i\bar{q}_l^i}\frac{x^i}{q_l^i} = \frac{\bar{q}_l^i}{\bar{x}^i}\frac{x^i}{q_l^i}.\square$$

The above proposition can be used to nonparametrically estimate the bounds to income elasticities using cross-country data. For the 5 country example data, one of the minimum subsistence consumption bundles is (1.06,0.08). For country 3, $\mathbf{p}^3 = 1, 1$, $\mathbf{q}^3 = 2.5, 2.5$,

$\bar{\mathbf{q}}^3 = 1.44, 2.42$, $x^i = 5$ and $\bar{x}^i = 3.86$. Hence $\varepsilon_1^3 = \frac{\bar{q}_1^3}{\bar{x}^3}\frac{x^3}{q_1^3} = \frac{1.44}{3.86}\frac{5}{2.5} = 0.75$ and similarly $\varepsilon_2^3 = 1.25$. For this particular $\boldsymbol{\gamma}$ we would therefore conclude that for country 3, good 1 is a necessity and good 2 is a luxury.

Rather than focusing on the income elasticities implied by a single element of $\mathcal{G}$, another approach is to calculate the range in which the income elasticities vary. For the example data, this information is presented in Table 5.2 along with descriptive statistics for the quantities and minimum subsistence consumption bundles (note that in this table, ε is calculated for the average country).

Table 5.2: Descriptive statistics for γ and ε, $N = 5$

	mean $\mathbf{q}$	min. $\mathbf{q}$	max. $\mathbf{q}$	min. γ	max. γ	min. ε	max. ε
good 1	3.65	1.25	7.00	0.00	1.14	0.89	1.05
good 2	3.05	0.50	7.50	0.00	0.49	0.89	1.25

5.5 Application to 1980 ICP Data

In this section, the bounds to the Ideal Afriat Marginal Index are constructed for the 42 countries in the 1980 ICP data set that were found to share common homothetic preferences (see Section 3.4 for details on how these countries were selected). The Ideal Afriat Index for the 42 countries which satisfied HARP, and ranking based on this index, are presented in Table 5.3 (these columns are replicated from Table 3.2).

For the 42 countries that satisfy HARP, we know $\mathcal{G}$ contains at least one vector, namely $\boldsymbol{\gamma} = \mathbf{0}$. The iterative process outlined in Section 5.3.5 was used to find 500 other vectors in $\mathcal{G}$ - this took 756 iterations. The descriptive statistics of the 500 minimum subsistence consumption bundles in $\mathcal{G}$ are presented in Table 5.4.

The impact of having countries with very low per capita consumption in the data set is evident; the upper bound to $\mathcal{G}$ shown in column 6 in Table 5.4 is constrained to be low (relative to the mean consumption for each good). A conclusion from this is that for the 1980 ICP data, conditional on affine homotheticity as a maintained hypothesis, the assumption of homotheticity is reasonable (except perhaps for food). Because the minimum subsistence bundles are relatively small, the bounds to the income elasticities of demand (calculated for the average country) shown in Table 5.4 are all close to one.

Despite this, we get some interesting results when we calculate the bounds to the Ideal Afriat Marginal Index (calculated over the 500 vectors in $\mathcal{G}$) which are shown in Table 5.3. The first thing to note is that the pairwise ranking between Cameroon and Senegal implied by the Ideal Afriat Index are contradicted by the bounds to the rankings implied by the Ideal Afriat Marginal Index. While $\mathbf{a}^*$ ranks Cameroon at 35th and Senegal at 36th, it is apparent that there also exists at least one other minimum subsistence consumption bundle which results in these pairwise rankings being reversed. The conclusion is that since there is no empirical reason to choose $\boldsymbol{\gamma} = \mathbf{0}$ (and hence use $\mathbf{a}^*$ as the welfare measure) over a $\boldsymbol{\gamma}$ which gives an alternative ranking of these countries, the countries should perhaps be ranked equivalent.[15]

Next, Table 5.3 can be used to classify countries according to Definition 5.6. In Figure

[15]Note, however, that $\mathbf{a}^*$ is only one of an infinite number of multilateral true welfare indexes that can be constructed in the homothetic case and it is equally possible that some of these other indexes may provide welfare rankings different to that implied by $\mathbf{a}^*$.

Table 5.3: Ideal Afriat Index and Ideal Afriat Marginal Index Bounds, 1980

			a^*_m		a^*_m *rank*					a^*_m		a^*_m *rank*	
	a*	rank	min.	max.	min.	max.		**a***	rank	min.	max.	min.	max.
Argentina	0.382	19	0.382	0.464	19	19	Japan	0.948	11	0.948	1.041	11	11
Austria	1.156	8	1.156	1.255	8	8	Kenya	-1.448	38	-1.633	-1.448	38	38
Belgium	1.262	5	1.262	1.364	5	5	Korea	-0.212	27	-0.212	-0.172	27	27
Bolivia	-0.527	29	-0.529	-0.516	29	29	Luxembourg	1.253	6	1.253	1.355	6	6
Botswana	-0.840	32	-0.876	-0.840	32	32	Madagascar	-1.357	37	-1.506	-1.357	37	37
Brazil	0.398	18	0.398	0.470	18	18	Malawi	-1.950	41	-2.354	-1.950	41	41
Cameroon	-1.148	35	-1.247	-1.148	35	36	Mali						
Canada	1.445	2	1.445	1.550	2	2	Morocco	-0.761	31	-0.782	-0.761	31	31
Chile	0.278	20	0.278	0.349	20	20	Netherlands						
Colombia	0.093	22	0.093	0.156	22	22	Nigeria						
Costa Rica	0.245	21	0.245	0.310	21	21	Norway	1.090	9	1.090	1.187	9	9
Ivory coast	-0.895	33	-0.939	-0.895	33	33	Pakistan						
Denmark	1.275	3	1.275	1.377	3	3	Panama	-0.076	24	-0.076	-0.030	24	24
Dominican Rep.							Paraguay						
Ecuador	-0.196	26	-0.196	-0.154	26	26	Peru	-0.084	25	-0.084	-0.033	25	25
El Salvador	-0.452	28	-0.456	-0.438	28	28	Philippines						
Ethiopia	-2.046	42	-2.598	-2.046	42	42	Poland						
Finland	0.933	12	0.933	1.024	12	12	Portugal	0.465	17	0.465	0.543	17	17
France	1.221	7	1.221	1.322	7	7	Senegal	-1.160	36	-1.254	-1.160	35	36
Germany FR	1.273	4	1.273	1.375	4	4	Spain						
Greece	0.574	13	0.574	0.658	13	13	Sri Lanka	-0.653	30	-0.666	-0.653	30	30
Guatemala	-0.054	23	-0.054	0.000	23	23	U.R. Tanzania						
Honduras							Tunisia						
Hong Kong	1.000	10	1.000	1.103	10	10	U.K.						
Hungary	0.490	16	0.490	0.568	16	16	U.S.	1.456	1	1.456	1.561	1	1
India	-1.646	39	-1.904	-1.646	39	39	Uruguay	0.540	14	0.540	0.625	14	14
Indonesia	-0.990	34	-1.050	-0.990	34	34	Venezuela	0.527	15	0.527	0.608	15	15
Ireland							Yugoslavia						
Israel							Zambia	-1.810	40	-2.137	-1.810	40	40
Italy							Zimbabwe						

Note: The minimum and maximum rankings based on a^*_m are *not* derived from the minimum and maximum of a^*_m (see text).

Table 5.4: γ and ε for 42 countries satisfying HARP, 1980

	mean q	min. q	max. q	min. γ	max. γ	min. ε	max. ε
1.1.1 Food	685.3	75.8	1322.1	0.0	75.5	0.92	1.01
1.1.2 Beverages	105.8	0.4	295.9	0.0	0.4	1.00	1.04
1.1.3 Tobacco	48.3	1.7	184.0	0.0	1.7	0.98	1.03
1.2.1 Clothing	183.8	10.8	855.0	0.0	10.8	0.96	1.04
1.2.2 Footwear	37.4	0.0	116.4	0.0	0.0	1.00	1.04
1.3.1 Gross rent	284.8	7.4	1072.1	0.0	7.4	0.99	1.03
1.3.2 Fuel	106.8	2.7	588.2	0.0	2.7	0.99	1.03
1.4.1 Furnishing etc.	132.1	1.4	591.3	0.0	1.4	1.00	1.04
1.4.2 HH goods etc.	97.0	7.6	245.0	0.0	7.6	0.94	1.03
1.5.1 Pvte. medical	200.1	3.1	758.8	0.0	3.1	0.99	1.04
1.6.1 Transport equip.	84.2	0.1	490.5	0.0	0.1	1.00	1.04
1.6.2 Trans. equip. (op.)	155.7	2.4	773.9	0.0	2.4	0.99	1.04
1.6.3 Purchased transport	53.9	0.7	260.3	0.0	0.7	1.00	1.04
1.6.4 Communication	35.8	0.1	181.9	0.0	0.1	1.00	1.04
1.7.1 Recreation	169.0	0.9	675.3	0.0	0.9	1.00	1.04
1.7.2 Education	191.7	14.1	503.9	0.0	14.0	0.95	1.03
1.8.1 Personal care	103.3	1.2	722.3	0.0	1.2	0.99	1.03
1.8.2 Other	224.4	0.0	923.5	0.0	0.0	1.00	1.04

Note: Note that the columns labeled min. γ and max. γ show the range over which γ varies - these columns themselves are not necessarily minimum subsistence bundles.

5.5, the Ideal Afriat Index and bounds to the Ideal Afriat Marginal Indexes are plotted for the 42 countries. Using Definition 5.6, 13 countries (or 31 percent of the sample) are classified as poor, 2 countries (El Salvador and Bolivia) cannot be classified, and the remaining are classified as rich.

5.5.1 Approximate multilateral true comparisons

As discussed in Chapter 4, the tests of HARP (and AHARP) are non-stochastic in that the data either pass the tests exactly, or else they don't and the hypothesis of utility maximisation is rejected. The Afriat efficiency index approach outlined in Section 4.3 can be used to conduct approximate multilateral true welfare comparisons. As shown in Section 4.4, 59 of the 60 countries in the 1980 ICP data set can be included in such a comparison; the Approximate Ideal Afriat Index (calculated for $e^* = 0.96$) and the implied rankings for these countries are shown in Table 5.5 (these numbers are replicated from Table 4.3).

For the 59 countries that satisfy HARP(0.96), we know $\mathcal{G}(0.96)$ (where the notation signifies that this set contains quantity vectors which are consistent with AHARP(0.96)) contains at least one vector, namely $\gamma = \mathbf{0}$. Once again, the iterative process was

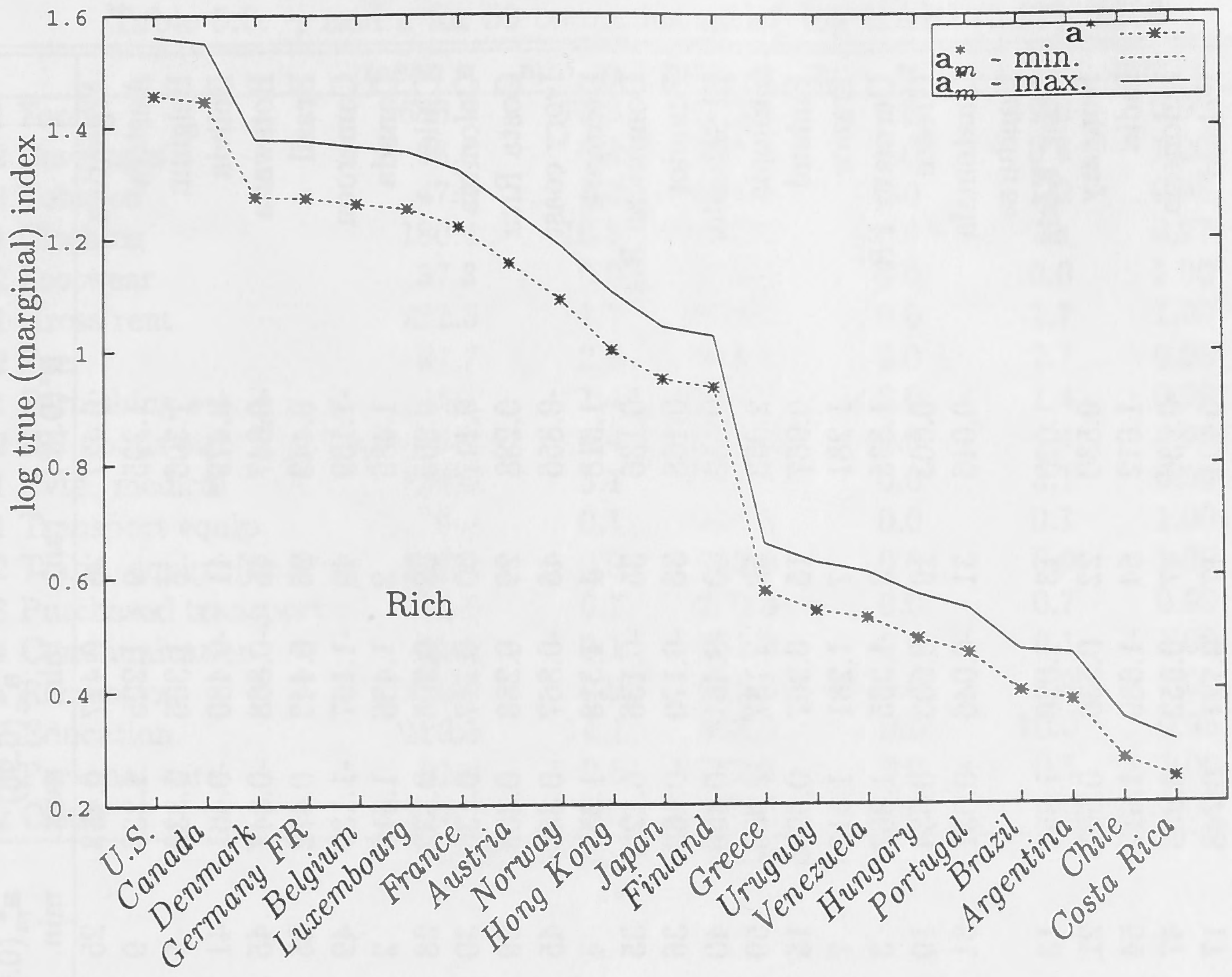

Figure 5.5: Ideal Afriat Index and Ideal Afriat Marginal Index bounds, 1980

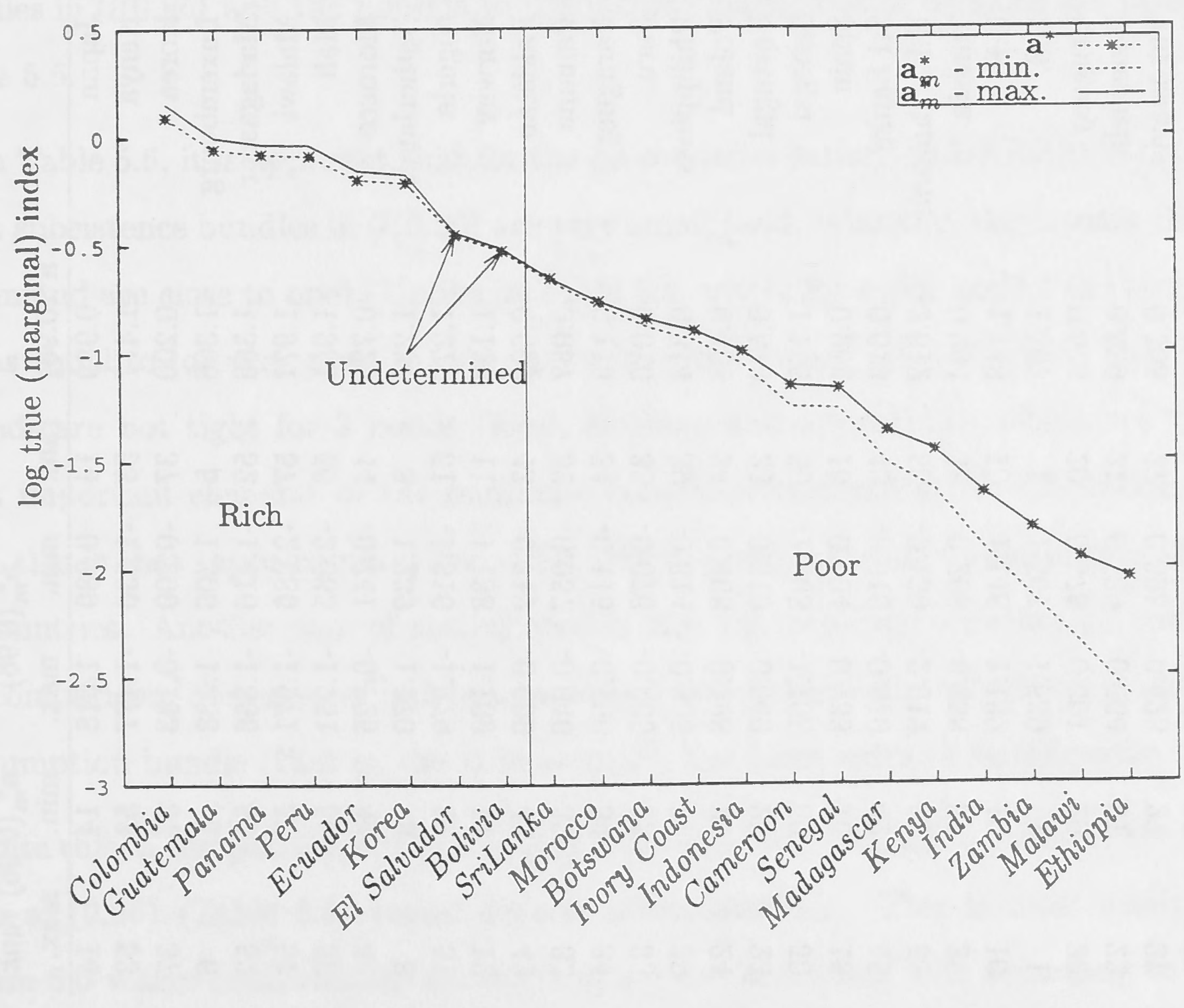

Figure 5.5: Ideal Afriat Index and Ideal Afriat Marginal Index bounds, 1980 (cont.)

Table 5.5: Approximate Ideal Afriat Index and Approximate Ideal Afriat Marginal Index Bounds, 1980

			$a^*_m(0.96)$		$a^*_m(0.96)$ *rank*					$a^*_m(0.96)$		$a^*_m(0.96)$ *rank*	
	$a^*(0.96)$	rank	min.	max.	min.	max.		$a^*(0.96)$	rank	min.	max.	min.	max.
Argentina	0.427	26	0.427	0.453	25	26	Japan	0.989	14	0.989	1.018	14	14
Austria	1.205	9	1.205	1.239	9	9	Kenya	-1.477	53	-1.536	-1.477	53	53
Belgium	1.306	6	1.306	1.338	5	6	Korea	-0.200	37	-0.200	-0.183	37	37
Bolivia	-0.479	41	-0.480	-0.461	41	41	Luxembourg	1.306	5	1.306	1.338	5	6
Botswana	-0.847	45	-0.868	-0.847	45	46	Madagascar	-1.366	52	-1.416	-1.366	52	52
Brazil	0.443	25	0.443	0.472	25	26	Malawi	-1.971	57	-2.086	-1.971	57	57
Cameroon	-1.139	49	-1.167	-1.139	49	49	Mali	-1.951	56	-2.065	-1.951	56	56
Canada	1.486	2	1.486	1.518	2	2	Morocco	-0.725	44	-0.741	-0.725	44	44
Chile	0.305	28	0.305	0.332	28	28	Netherlands	1.239	8	1.239	1.270	8	8
Colombia	0.117	30	0.117	0.136	30	30	Nigeria	-1.279	51	-1.316	-1.279	51	51
Costa Rica	0.288	29	0.288	0.308	29	29	Norway	1.138	11	1.138	1.169	11	12
Ivory coast	-0.856	46	-0.867	-0.856	45	46	Pakistan	-0.699	43	-0.712	-0.699	43	43
Denmark	1.318	4	1.318	1.351	4	4	Panama	-0.057	32	-0.057	-0.046	32	32
Dominican Rep.	-0.136	35	-0.136	-0.122	35	35	Paraguay	-0.115	34	-0.115	-0.100	34	34
Ecuador	-0.168	36	-0.170	-0.160	36	36	Peru	-0.076	33	-0.076	-0.065	33	33
El Salvador	-0.452	40	-0.461	-0.450	40	40	Philippines	-0.314	39	-0.314	-0.305	39	39
Ethiopia	-2.053	59	-2.187	-2.053	59	59	Poland	0.468	24	0.468	0.498	24	24
Finland	0.967	15	0.967	0.995	15	15	Portugal	0.519	23	0.519	0.549	23	23
France	1.281	7	1.281	1.313	7	7	Senegal	-1.165	50	-1.208	-1.165	50	50
Germany FR	1.335	3	1.335	1.367	3	3	Spain	0.904	16	0.904	0.933	16	16
Greece	0.603	19	0.603	0.634	19	19	Sri Lanka	-0.613	42	-0.616	-0.610	42	42
Guatemala	-0.046	31	-0.046	-0.031	31	31	U.R. Tanzania	-2.012	58	-2.139	-2.012	58	58
Honduras							Tunisia	-0.264	38	-0.268	-0.258	38	38
Hong Kong	1.020	13	1.020	1.052	13	13	U.K.	1.148	10	1.148	1.182	10	10
Hungary	0.530	22	0.530	0.561	21	22	U.S.	1.507	1	1.507	1.539	1	1
India	-1.612	54	-1.680	-1.612	54	54	Uruguay	0.575	20	0.575	0.601	20	20
Indonesia	-0.936	47	-0.952	-0.936	47	47	Venezuela	0.539	21	0.539	0.564	21	22
Ireland	0.721	17	0.721	0.748	17	17	Yugoslavia	0.395	27	0.395	0.423	27	27
Israel	0.693	18	0.693	0.726	18	18	Zambia	-1.768	55	-1.850	-1.768	55	55
Italy	1.138	12	1.138	1.171	11	12	Zimbabwe	-1.136	48	-1.163	-1.136	48	48

Note: See note to Table 5.3.

Table 5.6: γ and ε for 59 countries satisfying HARP(0.96), 1980

	mean **q**	min. **q**	max. **q**	min. γ	max. γ	min. ε	max. ε
1.1.1 Food	686.7	75.8	1505.0	0.0	19.6	0.98	1.00
1.1.2 Beverages	98.6	0.4	307.3	0.0	0.4	1.00	1.01
1.1.3 Tobacco	47.0	1.2	184.0	0.0	1.2	0.98	1.01
1.2.1 Clothing	180.4	10.8	855.0	0.0	7.1	0.97	1.01
1.2.2 Footwear	37.3	0.0	116.4	0.0	0.0	1.00	1.01
1.3.1 Gross rent	282.5	1.7	1072.1	0.0	1.7	1.00	1.01
1.3.2 Fuel	97.7	2.7	588.2	0.0	2.7	0.98	1.01
1.4.1 Furnishing etc.	126.1	1.4	591.3	0.0	1.4	0.99	1.01
1.4.2 HH goods etc.	88.0	4.4	245.0	0.0	4.4	0.96	1.01
1.5.1 Pvte. medical	192.0	3.1	758.8	0.0	3.1	0.99	1.01
1.6.1 Transport equip.	76.0	0.1	490.5	0.0	0.1	1.00	1.01
1.6.2 Trans. equip. (op.)	134.8	0.9	773.9	0.0	0.9	1.00	1.01
1.6.3 Purchased transport	50.6	0.7	260.3	0.0	0.7	0.99	1.01
1.6.4 Communication	30.2	0.1	181.9	0.0	0.1	1.00	1.01
1.7.1 Recreation	159.9	0.9	675.3	0.0	0.9	1.00	1.01
1.7.2 Education	182.5	14.1	503.9	0.0	10.0	0.95	1.01
1.8.1 Personal care	91.9	0.5	722.3	0.0	0.5	1.00	1.01
1.8.2 Other	217.7	0.0	923.5	0.0	0.0	1.00	1.01

Note: See note to Table 5.4.

used to collect an "adequate sample" of the vectors in $\mathcal{G}$ - 194 γs were collected, taking 10,000 iterations. The descriptive statistics of the 194 minimum subsistence consumption bundles in $\mathcal{G}(0.96)$ and the bounds to the income elasticities of demand are presented in Table 5.6.

From Table 5.6, it is apparent that for the 59 countries satisfying HARP(0.96), the minimum subsistence bundles in $\mathcal{G}(0.96)$ are very small (and, relatedly, the income elasticities of demand are close to one). Unlike in Table 5.4, where for every good l the largest γ_l in $\mathcal{G}$ was equal to (or very close to) the upper bound of $min_{i=1}^{N} q_l^i$, in Table 5.6 these upper bounds are not tight for 3 goods (food, clothing and education - which are the three most important elements of the minimum consumption bundle). A conclusion to draw from this is that (approximate) homotheticity is a very reasonable assumption for these 59 countries. Another way of stating this is that by including an extra 17 countries in the comparison, the range of indeterminateness associated with the minimum subsistence consumption bundle (that is, the volume of $\mathcal{G}$), has been reduced significantly.

Despite this, a comparison of the rankings from $\mathbf{a}^*(0.96)$ with the bounds to the rankings from $\mathbf{a}_m^*(0.96)$ (Table 5.5) reveal several inconsistencies. This is most easily seen in Figure 5.6 where countries are arranged along the horizontal axis according to $\mathbf{a}^*(0.96)$ and the bounds to the rankings from $\mathbf{a}_m^*(0.96)$ are plotted. It is apparent that for five pairs of countries, the pairwise ranking is dependent on what γ in $\mathcal{G}(0.96)$ is chosen for constructing the Ideal Afriat Marginal Index, and hence these countries should perhaps be ranked as equivalent.

In Figure 5.7, $\mathbf{a}^*(0.96)$ and the bounds to $\mathbf{a}^*_m(0.96)$ are plotted. Consistent with the above finding that $\mathcal{G}(0.96)$ is relatively small, the bounds to $\mathbf{a}^*_m(0.96)$ are very tight. In the figure, the classification of countries according to Definition 5.6 is also shown. Of the 59 countries, 17 (or 28.8 percent) are poor, 5 are neither rich nor poor, and the rest are rich.

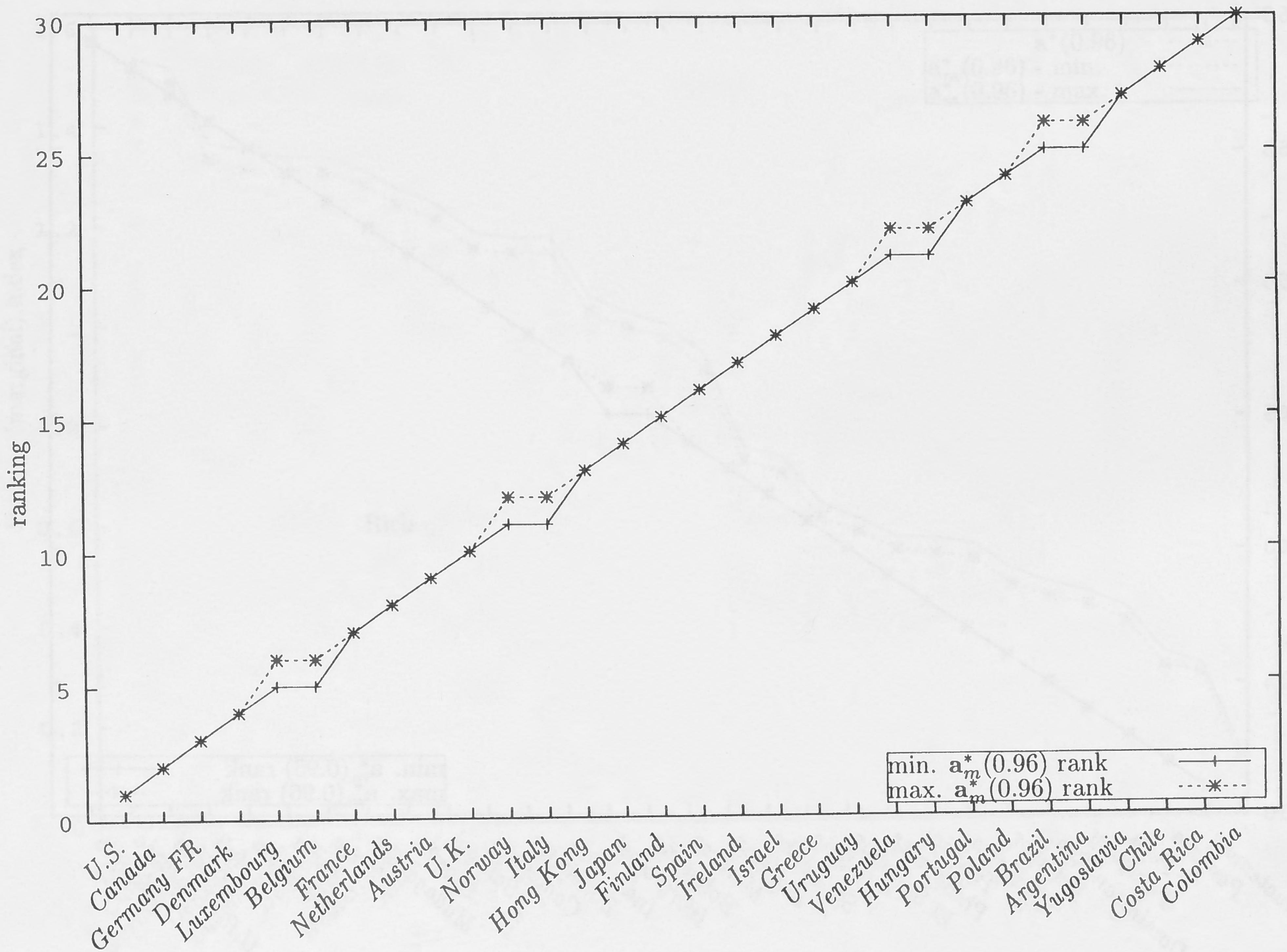

Figure 5.6: True welfare rankings implied by $\mathbf{a}^*(0.96)$ and $\mathbf{a}_m^*(0.96)$, 1980

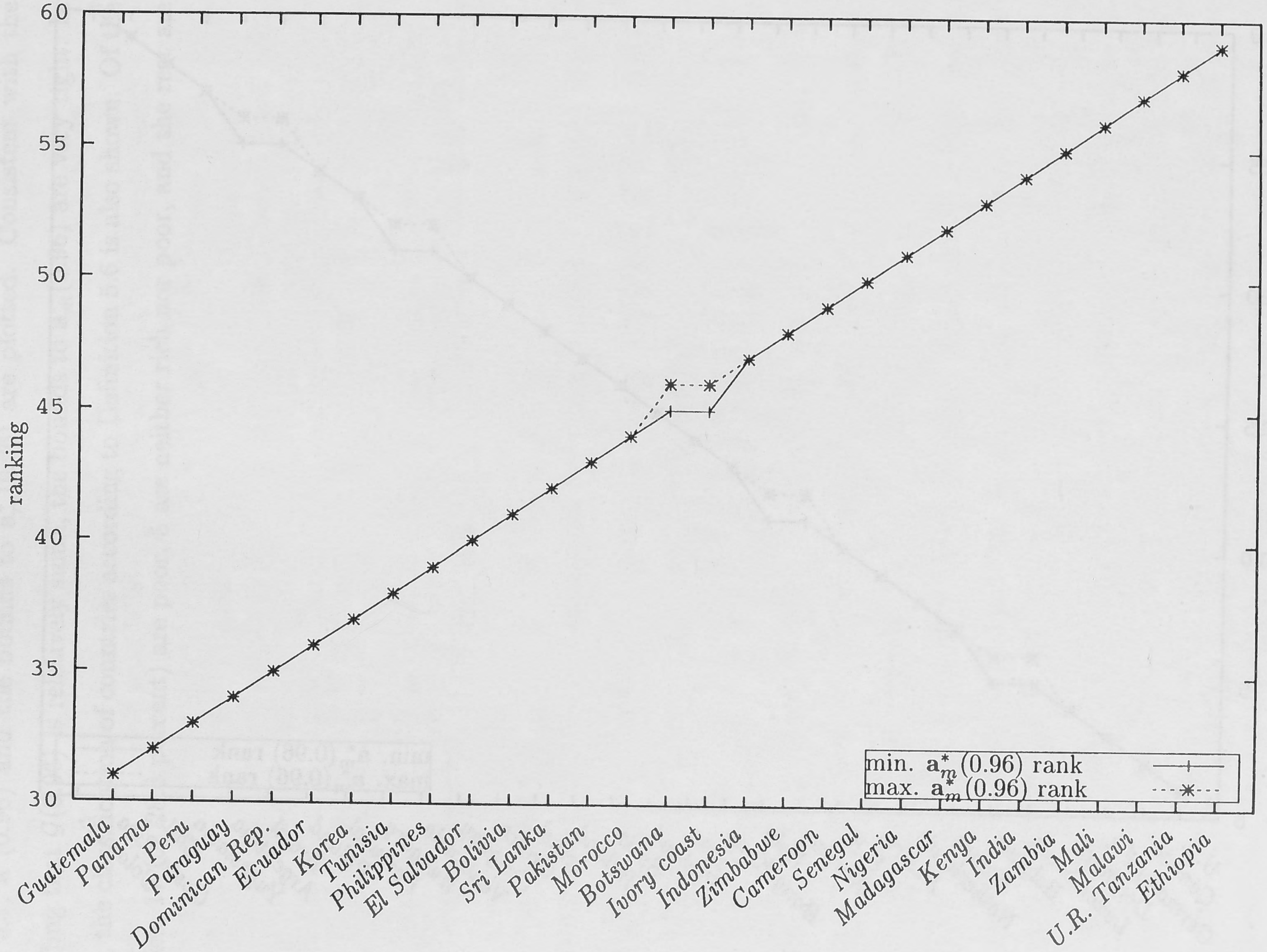

Figure 5.6: True welfare rankings implied by $\mathbf{a}^*(0.96)$ and $\mathbf{a}_m^*(0.96)$, 1980 (cont.)

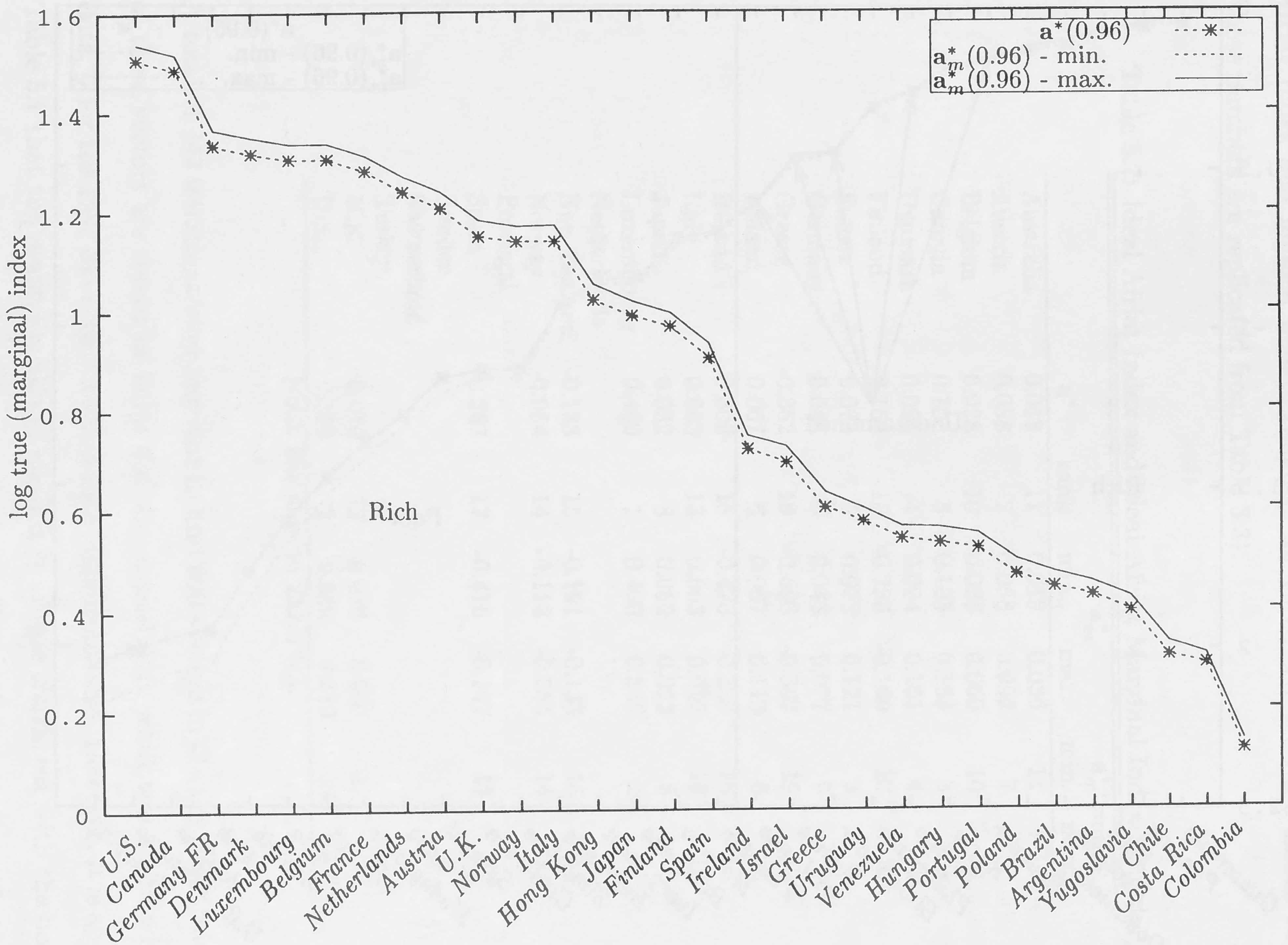

Figure 5.7: Approximate Ideal Afriat Index and Approximate Ideal Afriat Marginal Index bounds, 1980

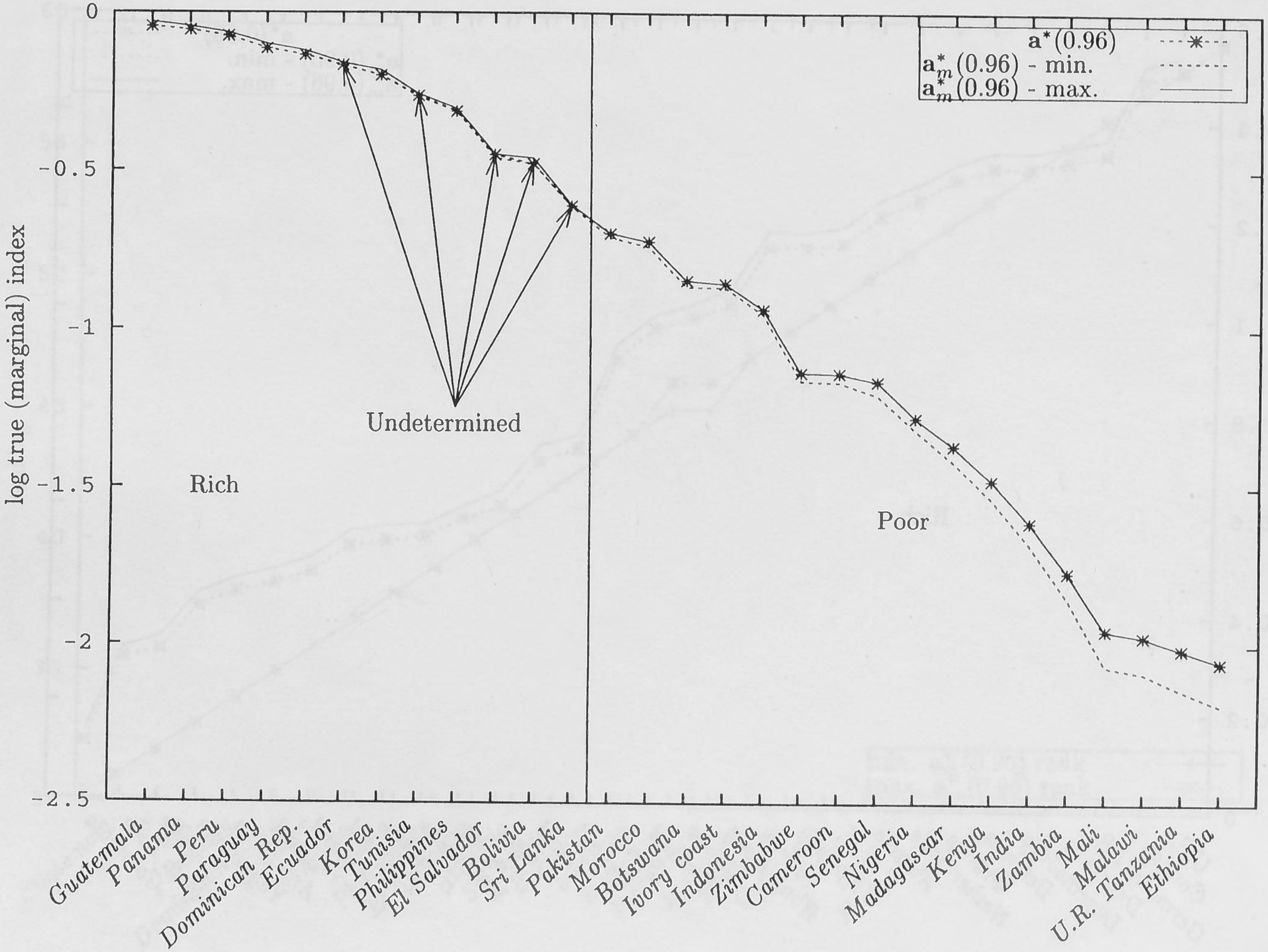

Figure 5.7: Approximate Ideal Afriat Index and Approximate Ideal Afriat Marginal Index bounds, 1980 (cont.)

5.6 Application to 1993 ICP Data

In this section, the above analysis is repeated for the 1993 ICP data on 24 OECD countries. In Section 3.5 it was found that 19 of the 24 countries satisfied HARP; the Ideal Afriat Index and implied welfare rankings for these countries are shown in Table 5.7 (these numbers are replicated from Table 3.3).

Table 5.7: Ideal Afriat Index and Ideal Afriat Marginal Index Bounds, 1993

	$\mathbf{a}^*$	rank	$\mathbf{a}_m^*$ min.	$\mathbf{a}_m^*$ max.	$\mathbf{a}_m^*$ *rank* min.	$\mathbf{a}_m^*$ *rank* max.
Australia	0.013	11	0.013	0.036	11	11
Austria	0.058	7	0.058	0.099	7	7
Belgium	0.028	10	0.028	0.060	10	10
Canada	0.105	3	0.105	0.168	3	3
Denmark	0.094	4	0.094	0.151	4	4
Finland	-0.160	16	-0.236	-0.160	16	16
France	0.073	5	0.073	0.121	5	5
Germany	0.043	9	0.043	0.077	9	9
Greece	-0.362	19	-0.596	-0.362	19	19
Iceland	0.067	6	0.067	0.113	6	6
Ireland	-0.300	18	-0.473	-0.300	18	18
Italy	0.003	13	0.003	0.017	13	13
Japan	0.052	8	0.052	0.083	8	8
Luxembourg	0.400	1	0.400	0.566	1	1
Netherlands						
New Zealand	-0.133	15	-0.191	-0.133	15	15
Norway	-0.084	14	-0.118	-0.084	14	14
Portugal						
Spain	-0.267	17	-0.416	-0.267	17	17
Sweden						
Switzerland						
Turkey						
U.K.	0.010	12	0.010	0.030	12	12
U.S.	0.359	2	0.359	0.513	2	2

Note: See note to Table 5.3.

A total of 742 iterations were required to find 500 vectors in $\mathcal{G}$ - the descriptive statistics of these vectors are shown in Table 5.8. Compared with what was found with the 1980 data, $\mathcal{G}$ for the 1993 data contains vectors of significant size. However, it is apparent from Table 5.7 that the country rankings implied by $\mathbf{a}^*$ are consistent with the bounds to the rankings associated with $\mathbf{a}_m^*$. Further, it should be noted that the bounds to the income elasticities of demand reported in Table 5.8 are too wide to be of any use empirically.

Table 5.8: γ and ε for 19 countries satisfying HARP, 1993

	mean q	min. q	max. q	min. γ	max. γ	min. ε	max. ε
1.1.1 Food	1774.4	1278.2	2747.8	0.0	1274.1	0.35	1.25
1.1.2 Beverages	318.2	189.1	460.5	0.0	188.8	0.49	1.42
1.1.3 Tobacco	256.7	146.2	1213.5	0.0	145.4	0.53	1.38
1.2.1 Clothing	608.0	330.7	1082.6	0.0	330.4	0.55	1.36
1.2.2 Footwear	125.8	53.3	260.9	0.0	52.8	0.67	1.35
1.3.1 Gross rent	1966.4	1112.7	2710.7	0.0	1101.2	0.60	1.29
1.3.2 Fuel	539.3	165.0	1422.7	0.0	164.5	0.82	1.36
1.4.1 Furnishing etc.	470.5	193.4	996.1	0.0	193.3	0.72	1.37
1.4.2 HH goods etc.	386.4	223.2	692.8	0.0	223.0	0.50	1.33
1.5.1 Pvte. medical	1114.3	228.6	2534.4	0.0	228.4	0.89	1.43
1.5.2 Public medical	678.0	0.0	1669.5	0.0	0.0	1.00	1.47
1.6.1 Transport equip.	476.6	146.9	1912.1	0.0	146.4	0.81	1.39
1.6.2 Trans. equip. (op.)	687.5	252.8	1428.0	0.0	252.3	0.76	1.37
1.6.3 Purchased transport	314.4	113.9	791.5	0.0	113.9	0.78	1.40
1.6.4 Communication	187.4	65.9	403.1	0.0	65.9	0.75	1.42
1.7.1 Recreation	908.1	314.8	1804.6	0.0	314.5	0.74	1.34
1.7.2 Education	1282.5	546.9	1932.9	0.0	545.5	0.69	1.39
1.8.1 Personal care	387.3	224.6	908.7	0.0	224.6	0.52	1.39
1.8.2 Other	1868.8	677.9	3191.9	0.0	677.8	0.77	1.33

Note: See note to Table 5.4.

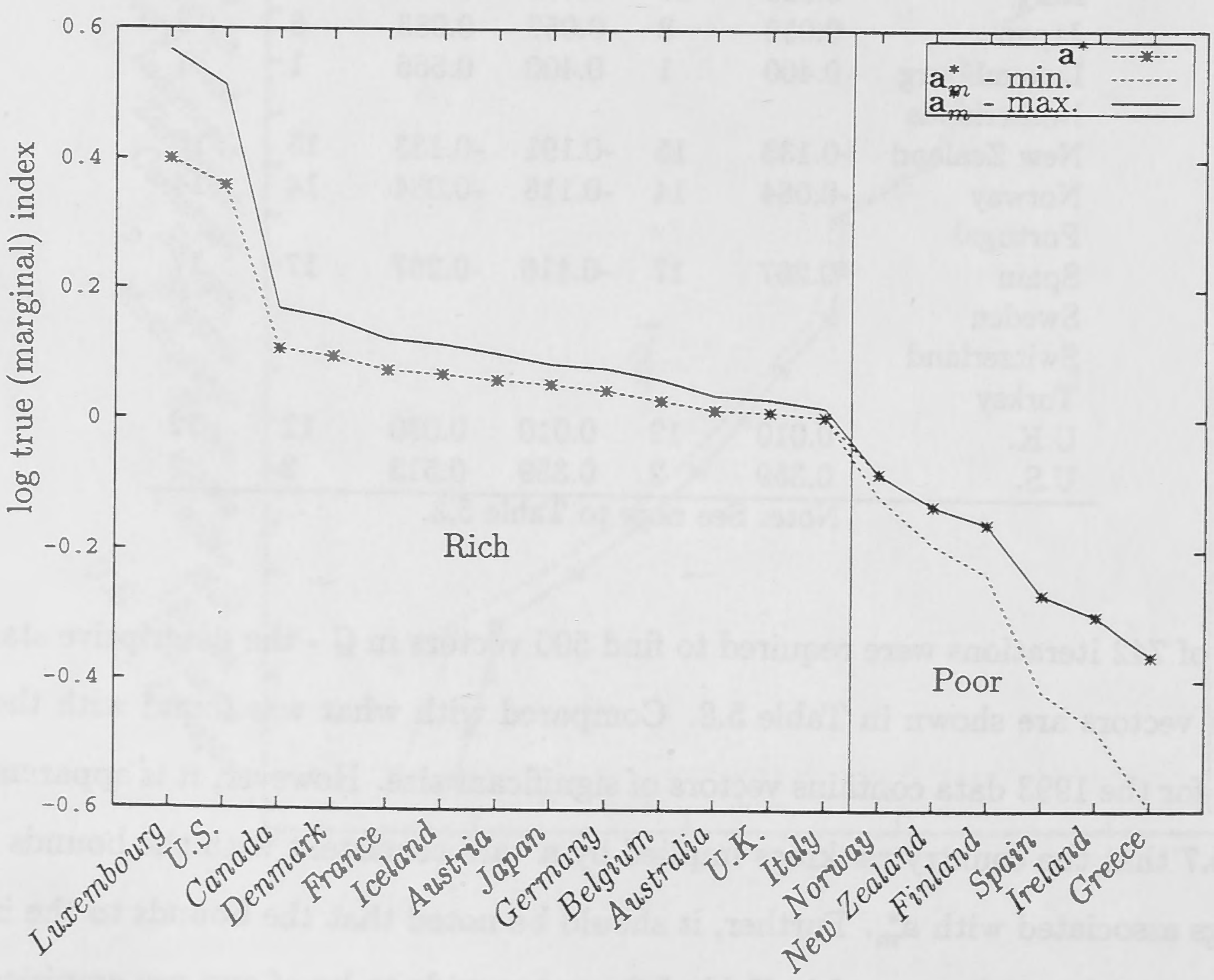

Figure 5.8: Ideal Afriat Index and Ideal Afriat Marginal Index bounds, 1993

Table 5.9: Approximate Ideal Afriat Index and Approximate Ideal Afriat Marginal Index Bounds, 1993

			$\mathbf{a}_m^*(0.98)$		$\mathbf{a}_m^*(0.98)$ *rank*	
	$\mathbf{a}^*(0.98)$	rank	min.	max.	min.	max.
Australia	0.067	13	0.067	0.108	12	14
Austria	0.115	8	0.115	0.170	8	8
Belgium	0.084	11	0.084	0.135	11	11
Canada	0.158	4	0.158	0.222	4	4
Denmark	0.135	5	0.135	0.194	5	5
Finland	-0.105	19	-0.118	-0.105	19	19
France	0.130	6	0.130	0.188	6	6
Germany	0.097	10	0.097	0.144	10	10
Greece	-0.308	23	-0.386	-0.308	23	23
Iceland	0.126	7	0.126	0.183	7	7
Ireland	-0.235	21	-0.282	-0.235	21	22
Italy	0.067	12	0.067	0.109	12	13
Japan	0.101	9	0.101	0.151	9	9
Luxembourg	0.440	1	0.440	0.557	1	1
Netherlands	0.029	15	0.029	0.060	15	15
New Zealand	-0.080	18	-0.085	-0.076	18	18
Norway	-0.013	17	-0.013	0.008	17	17
Portugal	-0.238	22	-0.289	-0.238	21	22
Spain	-0.211	20	-0.251	-0.211	20	20
Sweden	0.015	16	0.015	0.045	16	16
Switzerland	0.244	3	0.244	0.327	3	3
Turkey	-1.093	24	-1.829	-1.093	24	24
U.K.	0.066	14	0.066	0.109	12	14
U.S.	0.411	2	0.411	0.523	2	2

Note: See note to Table 5.3.

5.6.1 Approximate multilateral true comparisons

In Section 4.5 it was shown that approximate multilateral true comparisons could be made for all 24 countries in the 1993 ICP data set. The Approximate Ideal Afriat Index (calculated for $e^* = 0.98$) and the implied rankings are shown in Table 5.9 (these are taken from Table 4.4). It took 1741 iterations to collect 500 vectors which are elements of $\mathcal{G}(0.98)$; the descriptive statistics for these vectors are in Table 5.10. While the bounds to the elasticities of demand have been tightened compared with the bounds in Table 5.8, there are still generally quite wide (however, it appears that, as expected, food is a necessity).

As shown in Figure 5.9 the rankings from $\mathbf{a}^*(0.98)$ are not consistent with the bounds to the rankings from $\mathbf{a}_m^*(0.98)$ (Table 5.9). In particular, while the Ideal Afriat Index ranks Ireland at 21st and Portugal at 22nd, there exist γs in $\mathcal{G}(0.98)$ for which these rankings are reversed. The rankings between Australia, Italy and the U.K. are similarly

Table 5.10: γ and ε for 24 countries satisfying HARP(0.98), 1993

	mean **q**	min. **q**	max. **q**	min. γ	max. γ	min. ε	max. ε
1.1.1 Food	1737.0	1081.8	2747.8	0.0	1080.9	0.44	1.13
1.1.2 Beverages	309.2	13.4	480.6	0.0	13.4	1.00	1.24
1.1.3 Tobacco	233.9	84.4	1213.5	0.0	83.7	0.71	1.22
1.2.1 Clothing	596.3	271.8	1082.6	0.0	271.7	0.61	1.19
1.2.2 Footwear	127.5	53.3	260.9	0.0	53.2	0.64	1.21
1.3.1 Gross rent	1922.3	876.4	2710.7	0.0	632.5	0.84	1.20
1.3.2 Fuel	511.6	165.0	1422.7	0.0	164.9	0.75	1.20
1.4.1 Furnishing etc.	458.5	193.4	996.1	0.0	192.6	0.64	1.21
1.4.2 HH goods etc.	371.1	162.7	692.8	0.0	162.4	0.63	1.23
1.5.1 Pvte. medical	1087.0	165.4	2534.4	0.0	164.7	0.92	1.21
1.5.2 Public medical	620.3	0.0	1669.5	0.0	0.0	1.00	1.25
1.6.1 Transport equip.	436.2	49.5	1912.1	0.0	49.5	0.96	1.20
1.6.2 Trans. equip. (op.)	641.9	51.8	1428.0	0.0	51.7	1.00	1.25
1.6.3 Purchased transport	310.7	113.9	791.5	0.0	113.3	0.71	1.20
1.6.4 Communication	181.8	26.4	403.1	0.0	26.3	0.93	1.23
1.7.1 Recreation	865.8	79.9	1804.6	0.0	79.4	1.00	1.22
1.7.2 Education	1254.9	439.4	1932.9	0.0	433.6	0.74	1.22
1.8.1 Personal care	368.6	122.8	908.7	0.0	122.4	0.75	1.20
1.8.2 Other	1784.9	441.1	3191.9	0.0	439.7	0.85	1.18

Note: See note to Table 5.4.

indeterminate.

Finally, in Figure 5.10, $\mathbf{a}^*(0.98)$ and the bounds to $\mathbf{a}^*_m(0.98)$ are shown. Using Definition 5.6, 5 countries (or 20.8 percent of the sample) are classified as poor, and one country (New Zealand) cannot be classified as either poor or rich. Remember that this is a relative definition of poverty, that is, the five countries are only poor relative to the other OECD countries.

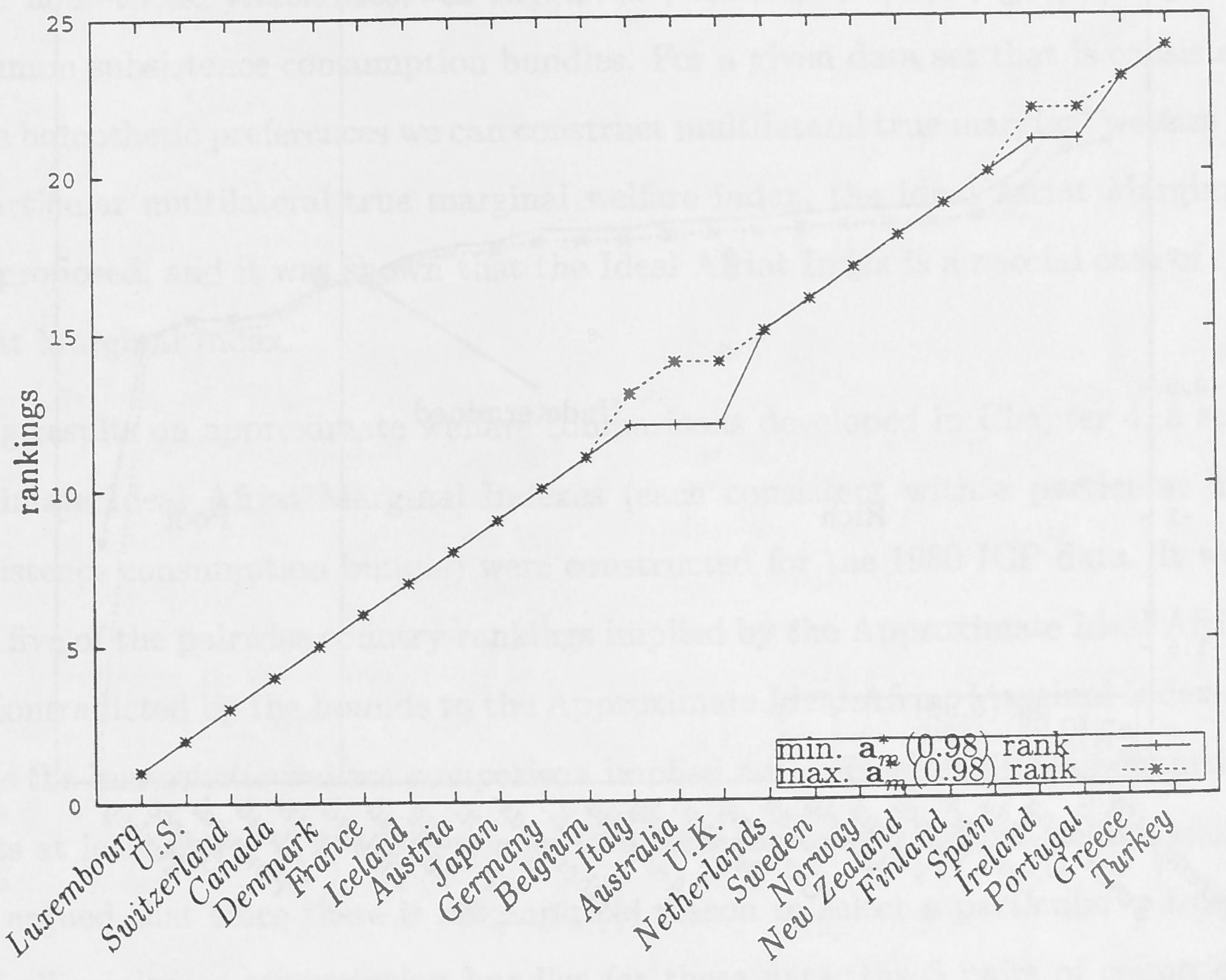

Figure 5.9: True welfare rankings implied by $\mathbf{a}^*(0.98)$ and $\mathbf{a}_m^*(0.98)$, 1993

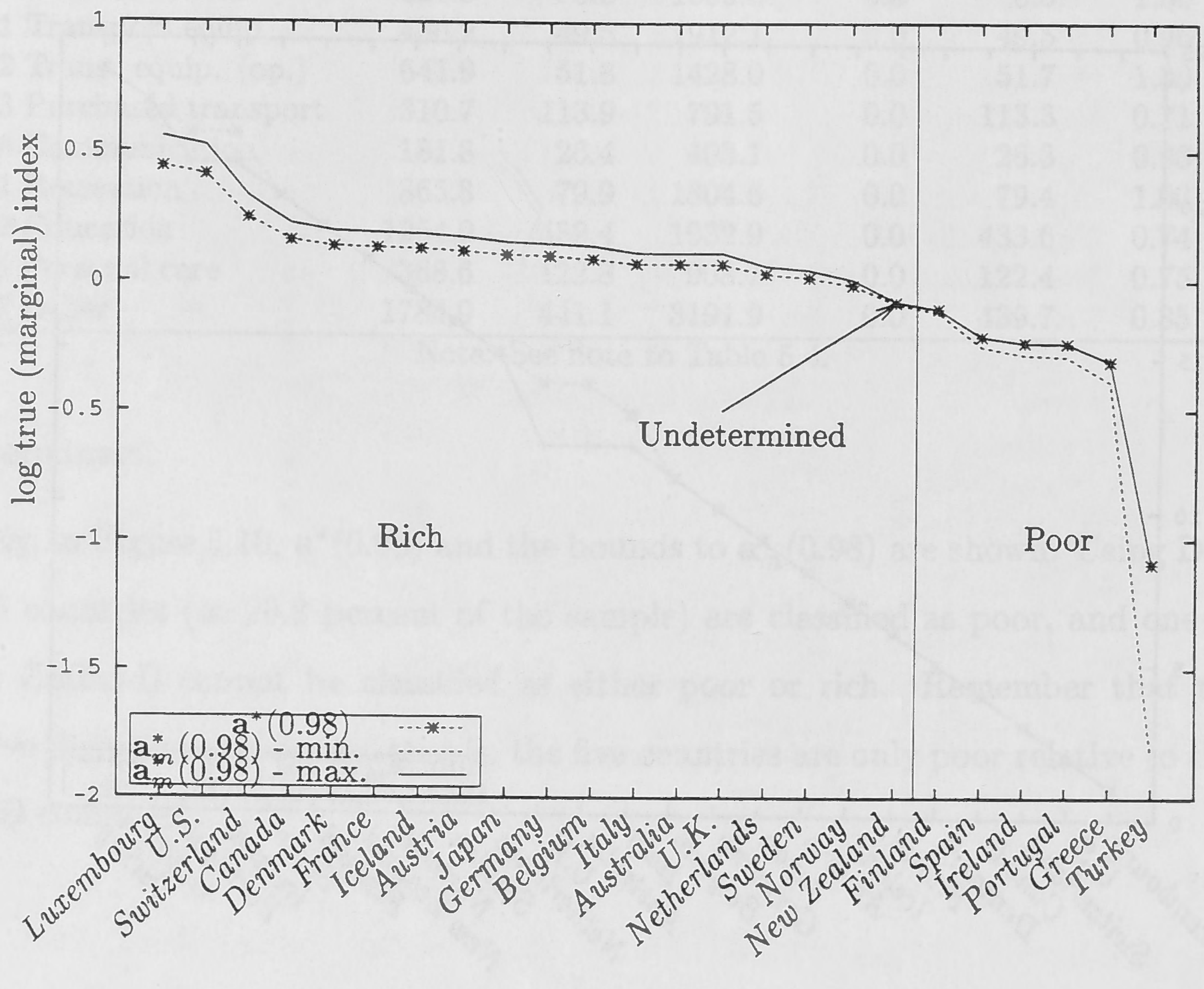

Figure 5.10: Approximate Ideal Afriat Index and Approximate Ideal Afriat Marginal Index bounds, 1993

5.7 Conclusions

In this chapter, a method for constructing multilateral true marginal welfare indexes was proposed and implemented using 1980 and 1993 ICP data. It was shown that a unique marginal welfare index exists when preferences are affine homothetic, that is, the expansion paths are linear and originate from a single point, the minimum subsistence consumption bundle. A method for testing a finite set of data for consistency with affine homothetic preferences was suggested - this involves testing for the existence of minimum subsistence consumption bundles. For a given data set that is consistent with affine homothetic preferences we can construct multilateral true marginal welfare indexes. A particular multilateral true marginal welfare index, the Ideal Afriat Marginal Index was proposed, and it was shown that the Ideal Afriat Index is a special case of the Ideal Afriat Marginal Index.

Using results on approximate welfare comparisons developed in Chapter 4, a set of Approximate Ideal Afriat Marginal Indexes (each consistent with a particular minimum subsistence consumption bundle) were constructed for the 1980 ICP data. It was found that five of the pairwise country rankings implied by the Approximate Ideal Afriat Index are contradicted by the bounds to the Approximate Ideal Afriat Marginal Indexes. Thus, while the homothetic welfare comparison implied a particular ranking of countries, there exists at least one $\gamma \neq \mathbf{0}$ which results in alternative true marginal welfare rankings. It was argued that since there is no empirical reason to select a particular γ from the set $\mathcal{G}$ of all minimum consumption bundles for these data, the 5 pairs of countries should perhaps be considered equivalent.

A method for ascertaining the poverty status of countries involving a comparison of the Ideal Afriat Index with the bounds to the Ideal Afriat Marginal Index was proposed. It was found that of the 59 countries for which (approximate) welfare comparisons can be made, 17 (or 28.8 percent) were poor.

With the 1993 ICP data for 24 OECD countries, several inconsistencies in the country rankings implied by the Ideal Afriat Index were also identified. It was further found that 5 countries (or 20.8 percent of the sample) were poor, although it should be emphasised that these countries are considered poor relative to the other OECD countries.

The methods presented in this chapter are applied in Chapter 6 to the construction of leisure-inclusive multilateral true welfare indexes.

Chapter 6

Leisure-Inclusive Welfare Comparisons

6.1 Introduction

The city businessman was at the pier of a small coastal village when a small boat with just one fisherman docked. Inside the small boat were several large yellowfin tuna. The businessman complimented the fisherman on the quality of his fish and asked how long it took to catch them. The fisherman replied only a little while. The businessman then asked why didn't he stay out longer and catch more fish? The fisherman said he had enough to support his family's immediate needs. The businessman then asked, but what do you do with the rest of your time? The fisherman said, "I sleep late, fish a little, play with my children, take a siesta with my wife, stroll into the village each evening where I sip wine and play guitar with my friends. I have a full and busy life, you see."

The businessman scoffed, "I have an MBA and could help you. You should spend more time fishing and with the proceeds buy a bigger boat, with the proceeds from the bigger boat you could buy several boats, eventually you would have a fleet of fishing boats. Instead of selling your catch to a middleman you would sell directly to the processor, eventually opening your own cannery. You would control the product, processing and distribution. You would need to leave this small coastal fishing village and move to the city where you will run your expanding empire."

The fisherman asked, "But how long will this all take?"

To which the businessman replied, "15-20 years."

"But what then", asked the fisherman?

The businessman laughed and said that's the best part. "When the time is right you would announce an IPO and sell your company stock to the public and become very rich, you would make millions."

"Millions? Then what?" asked the fisherman.

The businessman said, "Then you would retire. Move to a small coastal fishing village where you would sleep late, fish a little, play with your kids, take a siesta *with your wife, stroll to the village in the evenings where you could sip wine and play your guitar with your friends."*

Author unknown, via the internet.

Cross-country measures of economic welfare are generally based on per capita expenditures on goods and services. Such measures ignore an important commodity, leisure, which a consumer purchases implicitly by not working. In this chapter, the labour supply-goods demand framework used by Pencavel (1977), Cleeton (1982) and Riddell (1983), among others, in the construction of leisure-inclusive welfare measures is extended to a multilateral context, using the methods discussed in the last three chapters. In particular, a multilateral version of the index of real full income, the Leisure-Inclusive Ideal Afriat Marginal Index, is proposed and constructed using the 1993 ICP data on 24 OECD countries.

One of the drawbacks of estimating an index such as the index of real full income is that it requires an essentially arbitrary assumption about the amount of subsistence leisure time (that is, time spent on sleeping, eating and other necessary biological functions). However, the approach proposed in Chapter 5 for finding the bounds to the set of minimum subsistence bundles can be applied in the leisure-goods framework to nonparametrically estimate subsistence leisure time. The proposed method for estimating subsistence leisure is applied to the 1993 ICP data on 24 OECD countries and an upper-bound estimate of subsistence leisure is 7.7 hours/day.

This estimate of subsistence leisure is used in the construction of the Leisure-Inclusive Ideal Afriat Marginal Index, and the implied welfare rankings are markedly different to those found with the leisure-exclusive Ideal Afriat Index. For example, Japan has average hours of work per capita of almost twice the OECD average; this results in it falling from 9th in the leisure-exclusive welfare ranking to 19th when leisure time is included in the metric. While Luxembourg is ranked first using the Ideal Afriat Index, the fact that the average resident of the U.S. consumes more leisure time means that the U.S. is ranked first using the Leisure-Inclusive Ideal Afriat Marginal Index.

Some caveats need to be made about the methods and data used in these leisure-inclusive welfare comparisons. In particular, an indirect measure of leisure time consumed (calculated as the time available for market activities less hours of work) is used, and there are three reasons why we may not be accurately measuring the average amount of leisure time enjoyed in each country. First, there may be differences across countries in the amount of non-market work undertaken by the average resident. Second, since per capita hours of work are calculated as the product of average hours per worker and the employment-to-population ratio, differences in demographic composition across countries may affect the analysis (a country with a younger population will have higher recorded leisure, all other things equal). Finally, cross-country differences in labour market conditions will also affect average leisure time consumed, since unemployment is in effect measured as leisure time.

While these caveats are equally applicable to other research on measuring welfare in the leisure-goods framework (for example intertemporal studies within a particular country), it is to be expected that they might be particular important in a cross-country context. For this reason, the analysis was restricted to the 1993 OECD data; the OECD countries are by definition more homogenous, and therefore cross-country differences in non-market work and demographic composition are expected to be minimised. However, there are large differences in unemployment rates across the OECD countries, and a method is therefore proposed for "purging" the leisure variable for the effect of differences in labour market conditions.

The structure of the chapter is as follows. In Section 6.2, results on welfare measurement in the joint commodity demand-labour supply framework are presented, a multilateral version of the Allen real full income index, the Leisure-Inclusive Ideal Afriat Marginal Index, is proposed. Previous empirical work on leisure-inclusive welfare comparisons is reviewed in Section 6.3. In Section 6.4, leisure-inclusive welfare comparisons are conducted using the 1993 ICP data. Section 6.5 concludes the chapter.

6.2 Welfare Measurement in the Joint Commodity Demand-Labour Model

In this section, welfare measurement in the joint commodity demand-labour supply model is reviewed, and a multilateral true leisure-inclusive welfare index is proposed.

6.2.1 The joint commodity demand-labour supply model

In the leisure-augmented utility maximisation framework, the household is generally taken to be the decision unit. However, as with most studies in this area, the empirical analysis here is based on per capita aggregate data and it is therefore necessary to convert the household's decision problem into an equivalent problem that can be analysed with per capita data. As Barnett (1981, p.17) has noted, there are two problems associated with making this transition.

First, there is a problem associated with aggregating over household members; while all household members consume goods, not all members are in the labour force, and further, not all of those in the labour force need be employed. Barnett (1981) has explored the assumptions regarding preferences which are sufficient for the leisure-augmented utility maximisation framework to be expressed in terms of per capita demand functions which depend solely upon prices and per capita explanatory variables. For the same reasons as those outlined in Section 1.2.3, it is assumed here that such conditions are automatically met, and thus the model below is described in terms of a consumer-worker.

The second problem relates to the fact that in this study, leisure is constructed as the hours available for market activities, less the time spent working.[1] Since labour markets do not necessarily clear, there may exist a corner solution and per capita leisure consumption will not be on the per capita leisure demand function. The implication of this for leisure-inclusive welfare comparisons is that a country with high unemployment will record high per capita leisure (and thus be attributed a higher level of welfare) even

[1] Leisure is defined similarly in other studies of the joint commodity demand-labour supply framework (see, for example, Abbott and Ashenfelter (1976), Barnett (1981) and Coles and Harte-Chen (1985)). The reason for using an indirect measure of leisure is that comparable time-use survey data (which can provide a direct measure of leisure time) are not widely available; this is especially the case in a cross-country context.

though much of the measured leisure may be involuntary (that is, resulting from unemployment). Barnett (1981) suggests an approach for removing such problems which involves constructing a shadow price of leisure - this issue is discussed further below.

In the joint commodity demand-labour supply framework, the consumer-worker is assumed to maximise a well-behaved utility function $u(\mathbf{q}, l)$ with respect to a budget constraint $\mathbf{p}.\mathbf{q} = wh + m$, where $\mathbf{q}$ is a vector of K goods consumed at prices $\mathbf{p}$, l is the amount of leisure enjoyed, h is the hours worked at wage rate w, and m is exogenous non-labour income.[2] In the present study, all of the above variables are measured on a per annum basis. The individual's total endowment of time per year available for work (T) is calculated as the total number of hours in a year H ($52 \times 7 \times 24 = 8736$ hours) less an amount of time per year spent sleeping, eating and performing other biological functions (which can be thought of as "subsistence" leisure time). Available time is split between labour (h) and above-subsistence (or *supernumerary*) leisure (l), that is $T = h + l$. The two constraints are usually transformed into one: $\mathbf{p}.\mathbf{q} + wl = wT + m = y$, where y is called "full income" and is exogenous to the individual. In the standard model, the individual is assumed to be a price (wage) taker and to face no quantity constraints (although the latter assumption is discussed further below).

6.2.2 The bilateral true leisure-inclusive marginal welfare index

In this section, results on the construction of marginal indexes presented in Chapter 5 are use to define a leisure-inclusive marginal welfare index.

The Allen real full income index

There are three bilateral leisure-inclusive welfare index numbers which have been proposed in the literature; only one of these index numbers, the Allen index of real full income, is considered in this chapter.[3]

Define the full income expenditure function as:

$$f(\mathbf{p}, w, U^*) = \min_{(\mathbf{q},l)}\{\mathbf{p}.\mathbf{q} + wl : u(\mathbf{q}, l) \geq U^*\}.$$

[2] Taxation is ignored in the present study; see Baye and Black (1992) for an analysis of cost-of-living indexes in the presence of taxation and the goods-leisure decision.

[3] The other two index numbers (Allen real wage and Allen real non-labour income) are described in Riddell (1983).

With $\mathbf{p}$ and w fixed, this expenditure function is a monotonic increasing function of U and thus can be used in money metric welfare comparisons. The Allen real full income index comparing utility levels U^i and U^j at reference prices $\mathbf{p}^r$ and reference wages w^r is:

$$Y_{ij}^{A,r} = \frac{f(\mathbf{p}^r, w^r, u(\mathbf{q}^i, l^i))}{f(\mathbf{p}^r, w^r, u(\mathbf{q}^j, l^j))}.$$

The Allen real full income index thus gives the fraction of the cost of attaining country j's utility level required to attain country i's utility level at reference prices and wages. The natural candidates for $\mathbf{p}^r$ and w^r are $\mathbf{p}^j$, w^j and $\mathbf{p}^i$, w^i giving base-weighted and current-weighted true welfare indexes, which are known as the Laspeyres-Allen and Paasche-Allen full income indexes, respectively:

$$\begin{aligned} Y_{ij}^{LA} &= \frac{f(\mathbf{p}^j, w^j, u(\mathbf{q}^i, l^i))}{f(\mathbf{p}^j, w^j, u(\mathbf{q}^j, l^j))} = \frac{f(\mathbf{p}^j, w^j, u(\mathbf{q}^i, l^i))}{y^j} \\ Y_{ij}^{PA} &= \frac{f(\mathbf{p}^i, w^i, u(\mathbf{q}^i, l^i))}{f(\mathbf{p}^i, w^i, u(\mathbf{q}^j, l^j))} = \frac{y^i}{f(\mathbf{p}^i, w^i, u(\mathbf{q}^j, l^j))}. \end{aligned} \tag{6.1}$$

The Allen full income indexes Y_{ij}^{LA} and Y_{ij}^{PA} are not directly observable, however information on the *bounds* to the true indexes is provided by the Laspeyres and Paasche fixed-weight approximations:

$$Y_{ij}^{L} = \frac{\mathbf{p}^j\mathbf{q}^i + w^i l^i}{y^j}$$

$$Y_{ij}^{P} = \frac{y^i}{\mathbf{p}^i\mathbf{q}^j + w^i l^j}.$$

Analogous to Result 1.2 in Section 1.4.3, we can construct the following bounds:

$$Y_{ij}^{P} \leq Y_{ij}^{PA} \quad Y_{ij}^{LA} \leq Y_{ij}^{L}.$$

The Allen marginal real full income index

In Section 5.3.2, it was shown that if preferences are quasi homothetic, then there exist Allen marginal welfare indexes. It can similarly be shown that in the leisure-augmented utility maximisation framework, quasi-homothetic preferences leads to the existence of Allen marginal real full income indexes.

First, define the leisure-augmented price and quantity vectors for country i, as, respectively:

$$\begin{aligned}\tilde{\mathbf{p}}^i &= \{p_1^i, \dots, p_K^i, w^i\} \\ \tilde{\mathbf{q}}^i &= \{q_1^i, \dots, q_K^i, (H - h^i)\}.\end{aligned}$$

With quasi homotheticity, there exists minimum subsistence consumption bundles and hence the leisure-augmented quantity vector contains *total consumption*, that is, minimum subsistence consumption plus supernumerary consumption. Total consumption of leisure for country i is $H - h^i$, and this is split between subsistence leisure and supernumerary leisure l.

Following Section 5.3.1, with quasi homotheticity we have the Gorman Polar Form of the full income expenditure function:

$$f(U, \tilde{\mathbf{p}}) = a(\tilde{\mathbf{p}}) + b(\tilde{\mathbf{p}})U,$$

where $a(\tilde{\mathbf{p}})$ and $b(\tilde{\mathbf{p}})$ are positive and homogeneous of degree 1 functions of prices and U is leisure-inclusive utility above the subsistence level (normalised to be zero). The Allen marginal real full income index at references prices $\tilde{\mathbf{p}}^r$ is:

$$Y_{ij}^{AM,r} = \frac{f(U^i, \tilde{\mathbf{p}}^r) - a(\tilde{\mathbf{p}}^r)}{f(U^j, \tilde{\mathbf{p}}^r) - a(\tilde{\mathbf{p}}^r)}.$$

The Laspeyres-Allen and Paasche-Allen marginal real full income indexes are, respectively:

$$\begin{aligned}Y_{ij}^{LAM} &= \frac{f(U^i, \tilde{\mathbf{p}}^j) - a(\tilde{\mathbf{p}}^j)}{f(U^j, \tilde{\mathbf{p}}^j) - a(\tilde{\mathbf{p}}^j)} \\ Y_{ij}^{PAM} &= \frac{f(U^i, \tilde{\mathbf{p}}^i) - a(\tilde{\mathbf{p}}^i)}{f(U^j, \tilde{\mathbf{p}}^i) - a(\tilde{\mathbf{p}}^i)}.\end{aligned}$$

The indexes Y_{ij}^{LAM} and Y_{ij}^{PAM} are not directly observable. However, bounds to these indexes are provided by the Laspeyres and Paasche leisure-augmented marginal quantity indexes, defined respectively:

$$\begin{aligned}\tilde{Q}_{ij}^{LM} &= \frac{\tilde{\mathbf{p}}^j.(\tilde{\mathbf{q}}^i - \tilde{\gamma}^j)}{\tilde{\mathbf{p}}^j.(\tilde{\mathbf{q}}^j - \tilde{\gamma}^j)} \\ \tilde{Q}_{ij}^{PM} &= \frac{\tilde{\mathbf{p}}^i.(\tilde{\mathbf{q}}^i - \tilde{\gamma}^i)}{\tilde{\mathbf{p}}^i.(\tilde{\mathbf{q}}^j - \tilde{\gamma}^i)},\end{aligned}$$

where $\tilde{\gamma}^r$ is the leisure-inclusive minimum subsistence consumption bundle. Following the argument in Section 5.3.2, we can construct the following bounds:

$$\tilde{Q}_{ij}^{PM} \leq Y_{ij}^{PAM} \quad Y_{ij}^{LAM} \leq \tilde{Q}_{ij}^{LM}.$$

6.2.3 The Leisure-Inclusive Ideal Afriat Marginal Index

Analogous to Proposition 5.1, we can state that with affine homotheticity $\tilde{\gamma}^i = \tilde{\gamma}^j = \tilde{\gamma}$ and there exists a unique leisure-inclusive marginal index comparing the welfare of country i and j:

$$Y_{ij}^{LAM} = Y_{ij}^{PAM} = U^i/U^j.$$

Following Proposition 5.2, a test for the consistency of a given data set with affine homotheticity is that there exists a leisure-augmented quantity vector which can be defined as a leisure-augmented minimum subsistence consumption bundle $\tilde{\gamma}$. This test is carried out by selecting a particular leisure-augmented quantity vector (constrained so that none of the elements of the vector are negative and supernumerary consumption for all countries is non-negative) and constructing the appropriate matrix of logarithms of Laspeyres leisure-augmented marginal quantity indexes, $\{\tilde{L}_{ij}^m\} = \log \tilde{Q}_{ji}^{LM}$. Warshall's algorithm is then used to construct the minimum path matrix $\widetilde{\mathbf{M}}^m$, and if all of the diagonal elements of this matrix are non-negative, then the selected leisure-augmented quantity vector qualifies as a particular $\tilde{\gamma}$.

For a given $\tilde{\gamma}$, we can conduct leisure-inclusive marginal welfare comparisons using the leisure-inclusive version of the Ideal Afriat Marginal Index.

Definition 6.1. The Leisure-Inclusive Ideal Afriat Marginal Index: *For a given leisure-inclusive minimum subsistence consumption bundle $\tilde{\gamma}$, let $\tilde{\mathbf{a}}_m^+ \equiv (c^1, c^2, \ldots, c^N)$ and $\tilde{\mathbf{a}}_m^- \equiv (-r^1, -r^2, \ldots, -r^N)$ represent the vector of column means and the vector of negative row means, respectively, of the minimum path matrix defined over leisure-inclusive marginal quantities, $\widetilde{\mathbf{M}}^m$. Define the Leisure-Inclusive Ideal Afriat Marginal Index as $\tilde{\mathbf{a}}_m^* \equiv (\tilde{\mathbf{a}}_m^+ + \tilde{\mathbf{a}}_m^-)/2$, the vector of overall means.*

The properties and interpretation of the Leisure-Inclusive Ideal Afriat Marginal Index are analogous to that of the Ideal Afriat Index (see Proposition 3.4).

Analogous to Definition 5.5, we can define the set of leisure-inclusive minimum subsistence bundles $\tilde{\mathcal{G}}$. For a data set which is consistent with affine homotheticity, $\tilde{\mathcal{G}}$ will contain an infinite number of $\tilde{\gamma}$s, contained within a particular range. For a data set that is consistent with HARP, one of the vectors in $\tilde{\mathcal{G}}$ will be $\tilde{\gamma} = \mathbf{0}$. As discussed in Section 5.3.5, there is no empirical reason to base our leisure-inclusive welfare comparisons on a particular $\tilde{\gamma}$ in $\tilde{\mathcal{G}}$. However, in the empirical section below, an argument for using a particular $\tilde{\gamma}$ in $\tilde{\mathcal{G}}$ for constructing $\tilde{\mathbf{a}}_m^*$ is presented.

6.2.4 Non-market work, corner solutions and demographic composition

There are three caveats that need to be made about the above approach for conducting leisure-inclusive international comparisons of welfare. These caveats all relate to the degree to which the leisure variable employed in this study actually measures the leisure time enjoyed by the average consumer-worker in each country.

Non-market work

The definition of (supernumerary) leisure used in this study is such that it will include non-market work (for example, time spent cooking or looking after children). Thus, if there are large differences across countries in the time spent on such activities (which is to be expected, especially if countries from the entire development spectrum are included in the data set), this will not be reflected in the leisure-inclusive welfare measure. However, the potential bias arising from including non-market work time in leisure is minimised in this study by only constructing the leisure-inclusive welfare measures for the 24 OECD countries in the 1993 ICP data set. Since, by definition, these countries are more homogeneous, it is to be expected that cross-country differences in the average amount of time spent in non-market work will be minimised.

Corner solutions

It was mentioned above that a potential problem with using the leisure-augmented utility maximisation framework for conducting leisure-inclusive welfare comparisons arises if labour markets do not clear. For a country with a high unemployment rate, leisure time (as defined in this study) will include a component of "involuntary" leisure, i.e. unemployment. This is particularly a problem in the present context, where unemployment rates vary significantly across countries (for example, the unemployment rate for Japan in 1993 was 2.5 percent, compared with 22.7 percent in Spain).

Corner solutions can be removed if one uses a shadow price for leisure, as suggested by Barnett (1981). The shadow price for leisure will be equal to the wage rate if the labour market clears, and lower than the wage rate otherwise. However, the method for constructing a shadow wage rate suggested by Barnett (1981) involves estimating a parametric demand system, and this is at odds with the nonparametric methods which are the feature of the present study. In this study, therefore, the potential impact of corner solutions on the welfare measurement is minimised by adjusting the leisure variable to remove cross-country differences in unemployment rates.[4]

Demographic composition

The goods-leisure model used in this study is based on the existence of a consumer-worker, and consequently all variables are measured in per capita terms. However, the underlying theory involves the household as the relevant decision making unit, and one of the assumptions that is required to move to a consumer-worker model is that demographic variables do not influence consumption decisions. While this may be plausible for the consumption of goods (especially at the relatively broad aggregation used in this study), it is likely that the age-sex composition of a household will influence its consumption of leisure. Thus, by using per capita leisure in a cross-country leisure-inclusive welfare comparison, a country with a higher than average percentage of the population younger (or older) than the working age will be ranked higher simply because of demographic composition.

One could argue that consumption of leisure should be measured on a per adult equivalent basis rather than per capita basis.[5] However, previous work on leisure-inclusive welfare comparisons has tended to ignore the influence of age-sex composition on the consumption of leisure. One reason for this is because the use of an equivalence scale would introduce

[4]Future work will focus on whether it is possible to adapt the nonparametric test for rationing of Varian (1983, p.108) (and implemented by Swofford and Whitney (1994)) to the construction of true indexes in the presence of corner solutions.

[5]See Deaton and Muellbauer (1980) for details on the construction of equivalency scales.

arbitrariness into the analysis in terms of what weights are attached to the different age-sex categories. In this study, leisure is measured in per capita terms. However to minimise the possible bias arising from cross-country differences in demographic composition, the analysis will be restricted to the 1993 ICP data set on 24 OECD countries. It is to be expected that the differences in demographic composition across these countries is smaller than the differences across the 60 countries in the 1980 ICP data set.

6.3 Previous Work on Leisure-Inclusive Welfare Comparisons

Before applying the methods for leisure-inclusive welfare measurement to the ICP data, it is useful to briefly review previous research on leisure-inclusive welfare comparisons

6.3.1 Within-country welfare measurement

The majority of research on leisure-inclusive welfare comparisons has been concerned with temporal comparisons within a particular country. Pencavel (1977, 1979) proposed a constant-utility real wage rate which is the wage rate such that, given the prices and non-labour income for year t, the utility of the wage earner would be equal to that of a chosen base year.[6] Using parameters for the Stone-Geary utility function (estimated via the Linear Expenditure System (LES)), Pencavel found a 150 percent increase in the real wage between 1939 and 1967 in the U.S. This was more than double than that recorded by the Bureau of Labor Statistics (BLS) real spendable weekly earnings series and Pencavel concluded that the understatement of the BLS real spendable earnings series relative to his true real wage index arose because the BLS figures neglected the utility increasing effects both of shorter working hours and rising non-labour income.

A disadvantage of the functional approach is that the results of welfare measurement are dependent on the form of the utility function used. However, Coles and Harte-Chen (1985), using data identical to those used by Pencavel, compared constant-utility index numbers calculated from the LES and the Indirect Addilog System (IAS), and found that the constant-utility real wage rate was reasonably robust across the different specifications.

Empirical work on constructing real wage index numbers under the axiomatic approach has mainly focused on the Allen real wage index. Coles and Harte-Chen (1985) calculated the fixed-weight bounds to this index found that the functional real wage indexes based on the IAS and LES demand system frequently violated these bounds between 1939 and 1967 (however the majority of those violations occur during the Depression and war years). Kokoski (1987) constructed fixed-weight real wage indexes for different types of

[6] See, however, Lloyd (1979) for a criticism of this approach.

households using household-level data for the U.S. and found that some of the conclusions of Pencavel (1977) based on aggregate time series data may be misleading. In particular, Kokoski found that while there have been decreases in the national average of hours worked per week, increases in the labor force participation of women have led to a decline in leisure-consumption at the household-level. As result, it was concluded that leisure-inclusive welfare indexes, indicated smaller welfare increases (or larger decreases) over time than those indicated by a leisure-exclusive welfare index.

Swofford and Whitney (1987) and Patterson (1991) used nonparametric (revealed preference) methods to test the utility maximisation model when the utility function includes liquid assets, durables and leisure, in addition to consumption goods. The main objective of these articles was to test the assumption frequently used in applied demand research that the utility function exhibits weak separability between consumption goods and these "other" goods. Weak separability is very useful in consumer demand research since it implies a two-stage model for consumer behaviour; the demand for consumption goods can validly be modeled separately to the demand for other goods, thus allowing degrees of freedom in econometric estimation to be conserved. Applying tests developed by Varian (1983) to U.S. quarterly per capita data, Swofford and Whitney (1987) found that consumption goods and leisure meet the necessary and sufficient conditions for weak separability. In contrast, Patterson (1991) did not find any evidence of weak separability in U.K per capita data.

6.3.2 Cross-country welfare measurement

Dowrick and Quiggin (1993, 1994) developed a revealed preference approach to cross-country comparisons of welfare that explicitly includes leisure with consumption and investment. This method tests whether the representative consumer in country A could afford to buy the goods and services enjoyed by the average person in country B if consumer A worked B's hours, while receiving A's rate of pay for those labour hours. Dowrick and Quiggin (1994) found that in 1990, the average Japanese worked 23.7 hours per week, compared with the average Australian only working 16 hours per week. The authors therefore asked whether the consumption/investment bundle enjoyed by the average Australian is affordable to the representative Japanese if they were to reduce hours of work from 24 to 16 per week and it was found that adjusting for differences in leisure consumed using this approach markedly improved Australia's measured standard of living.

6.4 Application to 1993 ICP data

For the reasons discussed in Section 6.2.4, the method for constructing leisure-inclusive welfare comparisons was only applied to 1993 ICP data on 24 OECD countries (see Appendix A for details of the construction of the wage and leisure data).

In Figure 6.1, hourly wages (measured in ICP units of food) are plotted against average hours of work per week for the 24 OECD countries (both wages and hours of work are per capita estimates, not per worker estimates). There is a large amount of dispersion in average hours worked; the average citizens of Iceland and Japan both work over 20 hours per week (the average Japanese works 24.6 hours per week), while residents of Canada, Ireland and Spain work less than 12 hours per week on average.[6] One would expect leisure-inclusive welfare rankings to significantly differ from rankings which do not take into account differences in the consumption of leisure. In particular, countries such as Iceland and Japan should slip down the welfare rankings and Canada, Ireland and Spain should move up the country "league tables". Note, however, that the cross-country differences in hours worked largely reflect the fact that in 1993, unemployment rates varied markedly across the OECD countries; below a method for removing the effect of cross-country differences in unemployment rates from the leisure variable is proposed.

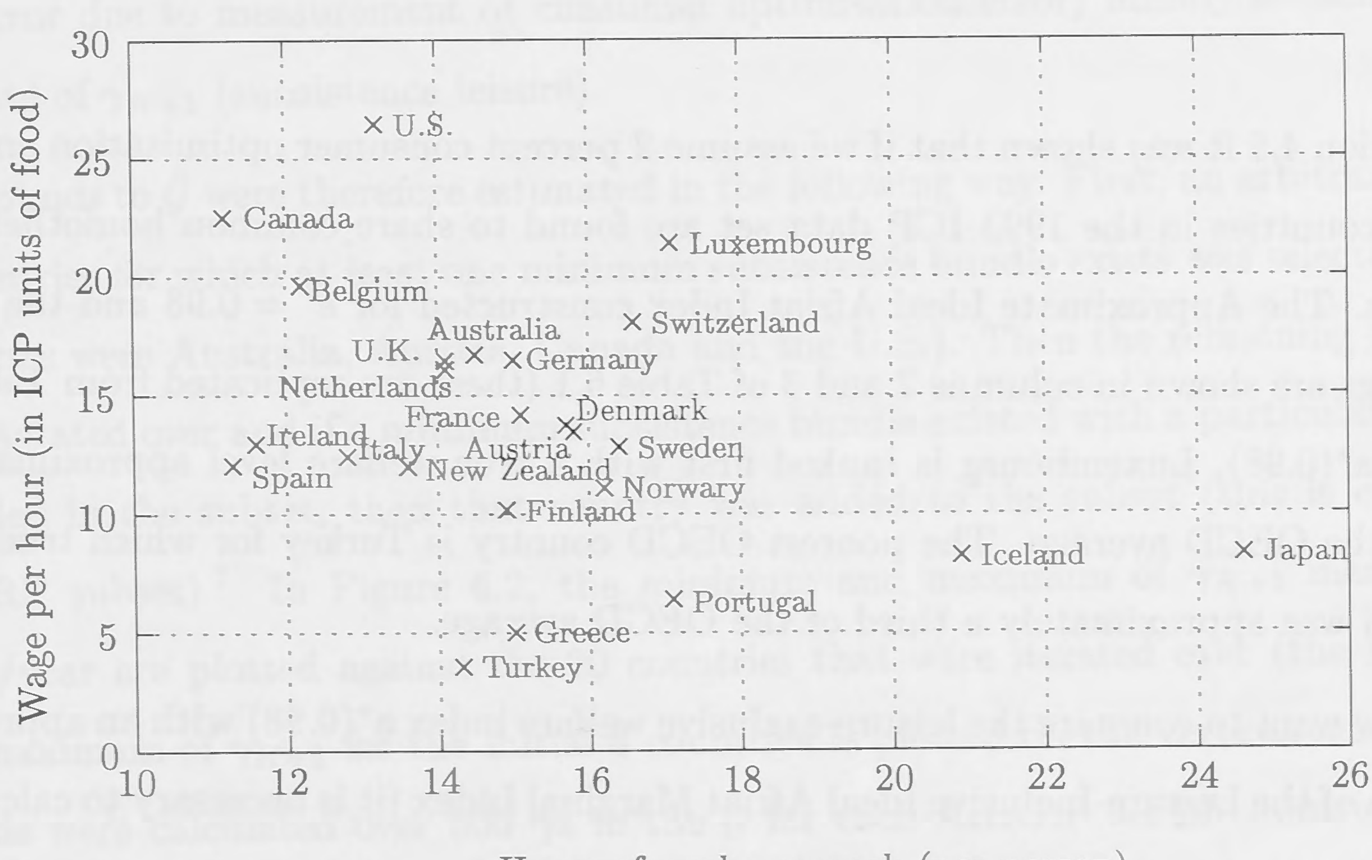

Figure 6.1: Hourly wages and hours of work, 1993

[6]Figure 6.1 indicates that average hours worked in the U.S. is low relatively to other OECD countries. This is partly explained by the fact that, as noted in Appendix A, hours of work data were only available for U.S. wage earners, not employees as a whole.

Table 6.1: Leisure-exclusive and leisure-inclusive welfare comparisons, 1993

			unadjusted data		*adjusted data*	
	$\mathbf{a}^*(0.98)$	rank	$\tilde{\mathbf{a}}_m^*(0.99)$	rank	$\tilde{\mathbf{a}}_m^*(0.99)$	rank
Australia	0.067	13	0.018	11	0.013	13
Austria	0.115	8	0.015	13	0.033	6
Belgium	0.084	11	0.040	4	0.043	5
Canada	0.158	4	0.058	2	0.054	3
Denmark	0.135	5	0.021	9	0.018	10
Finland	-0.105	19	-0.015	19	-0.042	20
France	0.130	6	0.023	7	0.016	12
Germany	0.097	10	0.019	10	0.024	9
Greece	-0.308	23	-0.046	21	-0.053	23
Iceland	0.126	7	-0.020	20	-0.005	16
Ireland	-0.235	21	-0.008	17	-0.031	18
Italy	0.067	12	0.029	5	0.027	7
Japan	0.101	9	-0.056	23	-0.036	19
Luxembourg	0.440	1	0.055	3	0.080	2
Netherlands	0.029	15	0.015	12	0.025	8
New Zealand	-0.080	18	0.001	14	-0.001	14
Norway	-0.013	17	-0.012	18	-0.003	15
Portugal	-0.238	22	-0.052	22	-0.045	21
Spain	-0.211	20	-0.001	15	-0.048	22
Sweden	0.015	16	-0.006	16	-0.007	17
Switzerland	0.244	3	0.027	6	0.046	4
Turkey	-1.093	24	-0.204	24	-0.215	24
U.K.	0.066	14	0.021	8	0.017	11
U.S.	0.411	2	0.077	1	0.089	1

Note: $\tilde{\mathbf{a}}_m^*(0.99)$ is calculated with $\gamma_{K+1} = 2394$ hours/year using the unadjusted leisure data and $\gamma_{K+1} = 2795$ using the adjusted leisure data.

In Section 4.5 it was shown that if we assume 2 percent consumer optimisation error then all 24 countries in the 1993 ICP data set are found to share common homothetic preferences. The Approximate Ideal Afriat Index constructed for $e^* = 0.98$ and the implied rankings are shown in columns 2 and 3 of Table 6.1 (these are replicated from Table 4.4). Using $\mathbf{a}^*(0.98)$, Luxembourg is ranked first with a true welfare level approximately 1.6 times the OECD average. The poorest OECD country is Turkey for which true welfare in 1993 was approximately a third of the OECD average.

We now want to compare the leisure-exclusive welfare index $\mathbf{a}^*(0.98)$ with an approximate version of the Leisure-Inclusive Ideal Afriat Marginal Index (it is necessary to calculate an approximate version of this index so as to include all 24 countries in the comparison). As discussed above, the general approach involves finding an "adequate sample" of bundles in $\tilde{\mathcal{G}}$ and calculating the bounds to $\tilde{\mathbf{a}}_m^*$, however it was decided to constrain $\tilde{\mathcal{G}}$ so that $\gamma_l = 0$ for all goods except leisure. Thus the minimum subsistence bundle only contains leisure; equivalently, preferences are assumed to be homothetic in all goods except leisure (and affine homothetic in leisure).

There were two reasons for constraining $\tilde{\mathcal{G}}$ in this manner. First, it is in keeping with other empirical studies in the goods-leisure framework which use supernumerary leisure in the analysis, but do not allow for the existence of other goods in the minimum subsistence consumption bundle. Second, by employing homotheticity for all goods except leisure it is easier to compare the welfare results from $\mathbf{a}^*$ and $\tilde{\mathbf{a}}^*_m$ i.e. it is only the inclusion of leisure which will be having an impact on welfare rankings.

The method for finding bundles in $\tilde{\mathcal{G}}$ used here is slightly different to that proposed in Chapter 5. In Table 5.10, descriptive statistics for 500 consumption bundles in $\mathcal{G}(0.98)$ were shown (these bundles qualify as minimum subsistence consumption bundles, as defined in Proposition 5.2, when $e^* = 0.98$). This approach was used in Chapter 5 because it enabled a comparison between the bounds to the Approximate Ideal Afriat Marginal Index constructed using these 500 vectors and $\mathbf{a}^*(0.98)$ (shown in Table 5.9). However, in this chapter we are more interested in finding an accurate estimate of minimum subsistence leisure and it doesn't necessarily make sense to be adding an error component to the data (which is what is happening when we use $e^* \neq 1$) prior to making this estimation. Further, using the Afriat efficiency index approach in the estimation of minimum subsistence leisure could lead to certain countries (whose consumption bundles may contain error due to measurement or consumer optimisation error) unduly influencing the estimate of γ_{K+1} (subsistence leisure).

The bounds to $\tilde{\mathcal{G}}$ were therefore estimated in the following way. First, an arbitrary subset of countries for which at least one minimum subsistence bundle exists was selected (these countries were Australia, Austria, Canada and the U.S.). Then the remaining countries were iterated over and if a minimum subsistence bundle existed with a particular country included in the subset, then that country was added to the subset (this is called the AHARP subset).[7] In Figure 6.2, the minimum and maximum of γ_{K+1} measured in hours/year are plotted against the 20 countries that were iterated over (the minimum and maximum of γ_{K+1} for the initial 4 countries is plotted on the vertical axis). These bounds were calculated over 500 γs in the $\tilde{\mathcal{G}}$ for each AHARP set of countries. If the inclusion of a particular country into the subset resulted in a minimum subsistence bundle

[7] A major difference with this process compared to the method used in Chapter 5 is that it is possible for $\tilde{\mathcal{G}}$ to not contain $\tilde{\gamma} = \mathbf{0}$.

not existing, then there is no marker for that country (such countries included Belgium, Iceland and Ireland).

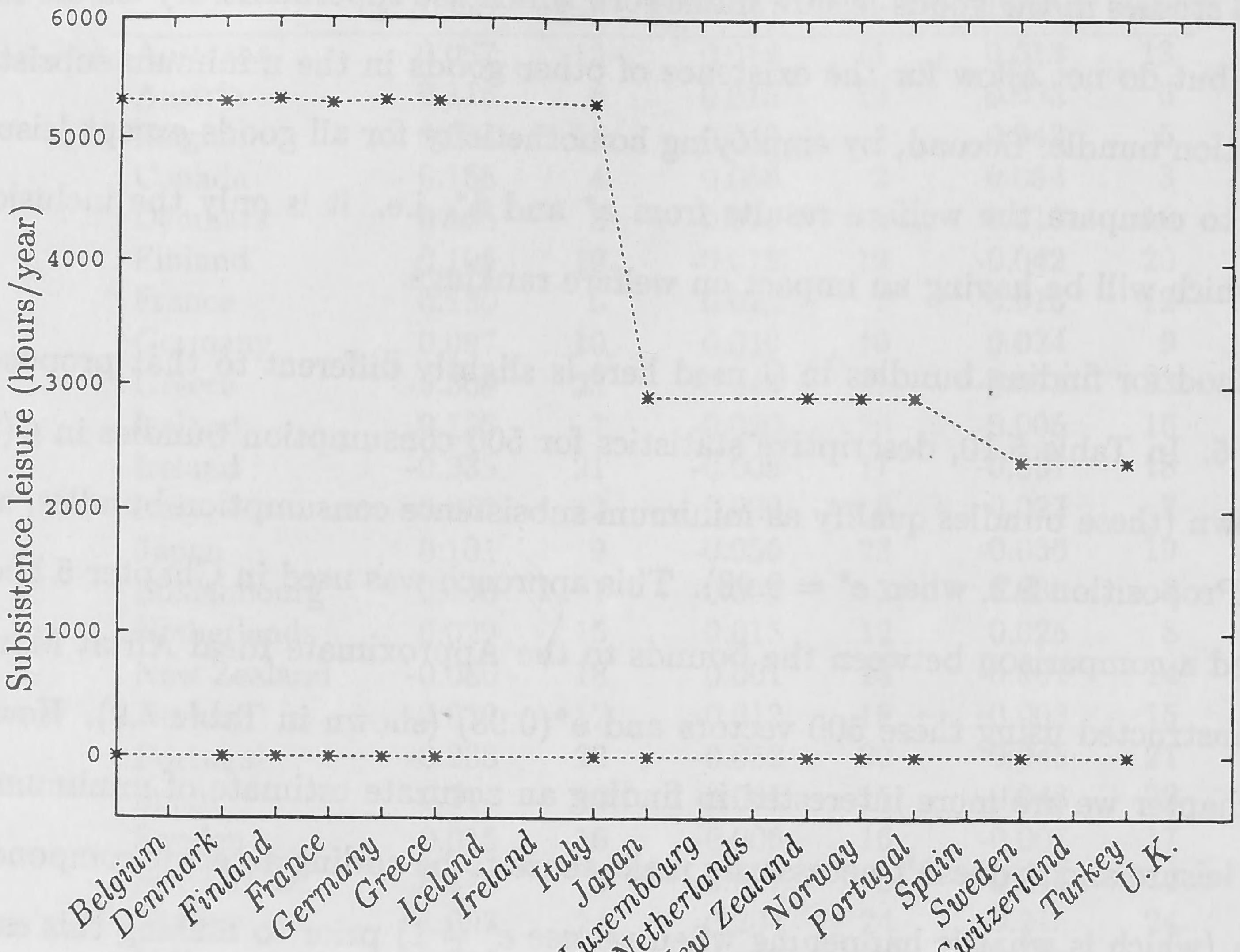

Figure 6.2: Estimated bounds to subsistence leisure, 1993

The definition of the minimum subsistence bundle in Proposition 5.2 implies that as more countries are added to the AHARP subset the bounds to $\tilde{\mathcal{G}}$ cannot be getting wider.[8] From Figure 6.2, it is apparent that the estimated upper bound to γ_{K+1} was tightened as Japan and Sweden were added to the AHARP set. The upper bound was reduced by around 50 percent (from 5295 to 2915 hours/year) when Japan was added, and thus it is possible to identify Japan as an "outlier" country. However, in the absence of a formal empirical method for identifying outlier countries, Japan was kept in the final AHARP subset of 16 countries.[9] For this subset of countries, minimum subsistence leisure was 0 hours/year, maximum γ_{K+1} was 2394 hours/year and mean γ_{K+1} was 1235 hours/year). The maximum estimate of subsistence leisure is therefore approximately 6.7 hours per day. While there is no empirical reason to base the leisure-inclusive welfare comparisons

[8]Note, however, that since we are only taking a sample of the vectors in $\tilde{\mathcal{G}}$ it is possible that the *estimated* bounds to this set could increase as more countries are added to the AHARP set.

[9]It took 1527 iterations to collect 500 vectors in $\tilde{\mathcal{G}}$ for these 16 countries.

on a particular γ in $\tilde{\mathcal{G}}$, commonsense tells us that the maximum estimate of 6.7 hours per day is a reasonable choice. However, this nonparametric estimate of minimum subsistence leisure is around half the estimate of 4680 hours/year (12.9 hours/day) which was used by Kuznets (1952) in the construction of a leisure-inclusive GDP series for the U.S.).[10]

The Approximate Leisure-Inclusive Ideal Afriat Marginal Index was calculated using $\gamma_{K+1} = 2394$ hours/year (see columns 4 and 5 of Table 6.1). It was necessary to assume consumer optimisation error of approximately 1 percent ($e^* = 0.99$) in order to include all 24 countries in the welfare comparison. It is apparent that the large differences in average hours worked (and hence leisure time consumed) identified in Figure 6.1 lead to radically different conclusions about the relative welfare levels of OECD countries. The fact that the average citizen of the U.S. consumes more leisure than the average resident of Luxembourg results in the U.S. being ranked first in the leisure-inclusive welfare comparison (compared with being ranked second according to $\mathbf{a}^*(0.98)$). Some of the other changes in rankings are even more marked. The average Japanese works almost twice as hard as the average resident of OECD countries and this results in Japan being ranked second last on the basis of $\tilde{\mathbf{a}}^*_m(0.99)$ compared with being ranked ninth using $\mathbf{a}^*(0.98)$ (and Iceland slips from 7th to 20th). The differences in rankings produced by $\mathbf{a}^*(0.98)$ and $\tilde{\mathbf{a}}^*_m(0.99)$ is evident in a Pearson's correlation coefficient of 0.724.

Since all countries consume leisure (and leisure is a relatively large component of full income), the inclusion of leisure into the welfare metric results in welfare being more equitably distributed across the countries. For example, the true leisure-inclusive welfare level of the U.S. is only approximately 8 percent above the OECD average, while that of Turkey is approximately 82 percent of the mean.

6.4.1 Leisure data adjusted for differences in unemployment rates

The fact that per capita hours worked is indirectly measured raises questions about the validity of leisure-inclusive welfare comparisons in Table 6.1. In particular, one has to question whether the average Japanese really works almost 10 hours per week more than the average resident of OECD countries, and consequently the accuracy of the finding in

[10] Kuznets (1952, p.64) estimated that the total number of hours per week available for work is 78 hours - and thus $T = 78 \times 52 = 4056$ hours/year. Hence Kuznets' implied estimate of subsistence leisure is $\gamma_{K+1} = H - T = 8736 - 4056 = 4680$ hours/year. Kuznets' estimate of T has been used by authors such as Barnett (1981).

Table 6.1 that Japan falls from ninth on a leisure-exclusive welfare ranking to 19th when leisure is included in the metric.

As argued above, the OECD countries have relatively similar demographic structures, and thus cross-country differences in the employment-to-population ratio are less likely to be due to demographic factors. Evidence for this is shown in Figure 6.3, where there does not appear to be a relationship between the percentage of the population of working age and the employment-to-population ratio (e/p). When the percentage of the population of working age was regressed against e/p also, the estimated slope coefficient was not significantly different from zero, thus giving further support to this observation.[11]

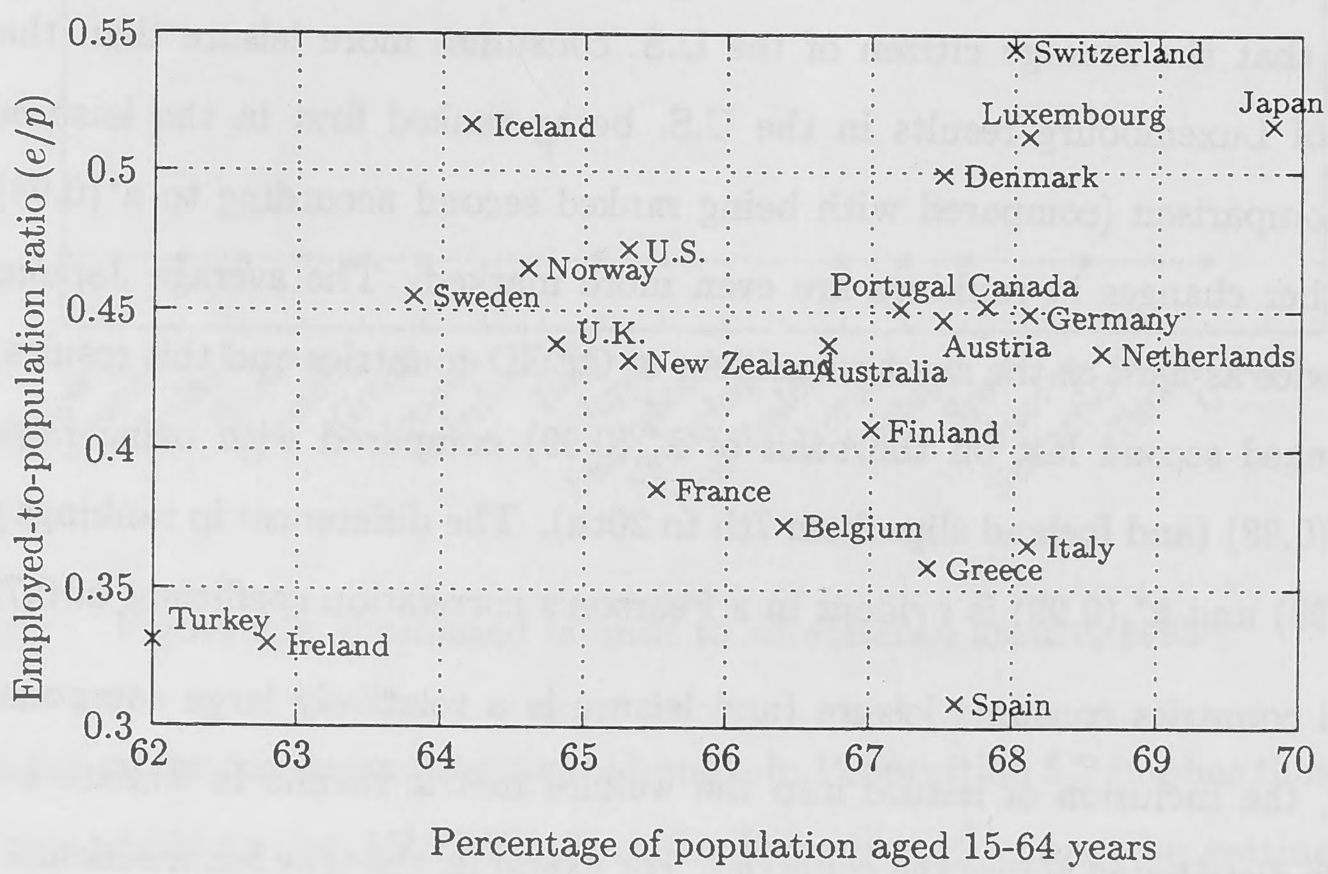

Figure 6.3: Employed-to-population ratio, percentage of population aged 15-64 years, 1993

It is to be expected, however, that labour market conditions in each country will be influencing both the employment-to-population ratio (which can be expected to rise during periods of labour market growth) and average hours worked by workers (which will similarly be positively influenced by a strong labour market). The labour market conditions in each country can be proxied by the unemployment rate, and from Figure 6.4 there indeed appears to be a negative relationship between the unemployment rate and average hours of work per week per capita.

[11] In contrast, when this exercise was performed for the 60 countries in the 1980 ICP data set, there was evidence to suggest that cross-country differences in demographic composition were influencing the employment-to-population ratio.

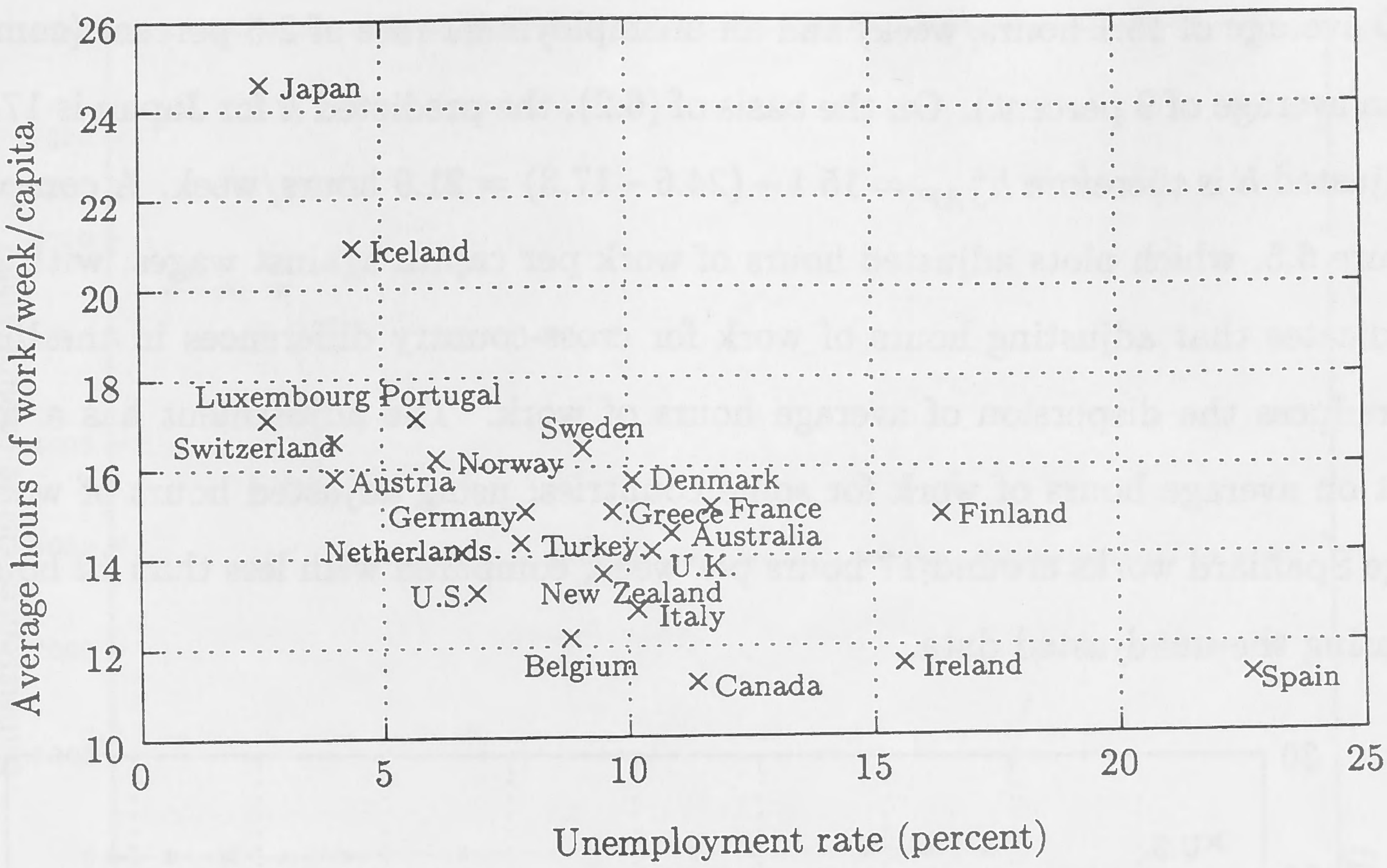

Figure 6.4: Unemployment rate, average hours of work/week/capita, 1993

The apparent negative correlation between the unemployment rate and average hours of work per capita is confirmed in the following regression results:[12]

$$h_i = \underset{(1.038)}{18.864} - \underset{(0.103)}{0.416}UR_i$$

$$R^2 = 0.651; \quad N = 24, \tag{6.2}$$

where UR_i is the unemployment rate in country i. The regression indicates that average hours of work per capita are indeed higher in countries with lower rates of unemployment. The regression coefficients can be used to correct hours of work per capita (and hence the measure of leisure time used in the welfare comparisons) for cross-country differences in the unemployment rate. In particular, the estimated parameters in (6.2) were used to find hours of work per capita $\hat{h}$ for each country; the residual for country i, $h_i^r = h_i - \hat{h}_i$, is then used to construct an adjusted hours of work variable, h_i^a which is free of the influence of the business cycle: $h_i^a = \overline{h} + h_i^r$, where $\overline{h}$ is the mean hours of work per capita.

To illustrate this adjustment for the business cycle, consider the example of Japan which in 1993 had average hours of work per capita of 24.6 hours/week (compared with the

[12] From Figure 6.4, a non-linear functional form should perhaps be used in this regression.

OECD average of 15.1 hours/week) and an unemployment rate of 2.5 percent (compared with an average of 9 percent). On the basis of (6.2), the predicted h for Japan is 17.8 and the adjusted h is therefore $h^a_{JAP} = 15.1 + (24.6 - 17.8) = 21.9$ hours/week. A comparison of Figure 6.5, which plots adjusted hours of work per capita against wages, with Figure 6.1 indicates that adjusting hours of work for cross-country differences in the business cycle reduces the dispersion of average hours of work. The adjustment has a marked impact on average hours of work for some countries; using adjusted hours of work, the average Spaniard works around 17 hours per week, compared with less than 12 hours per week using the unadjusted data.

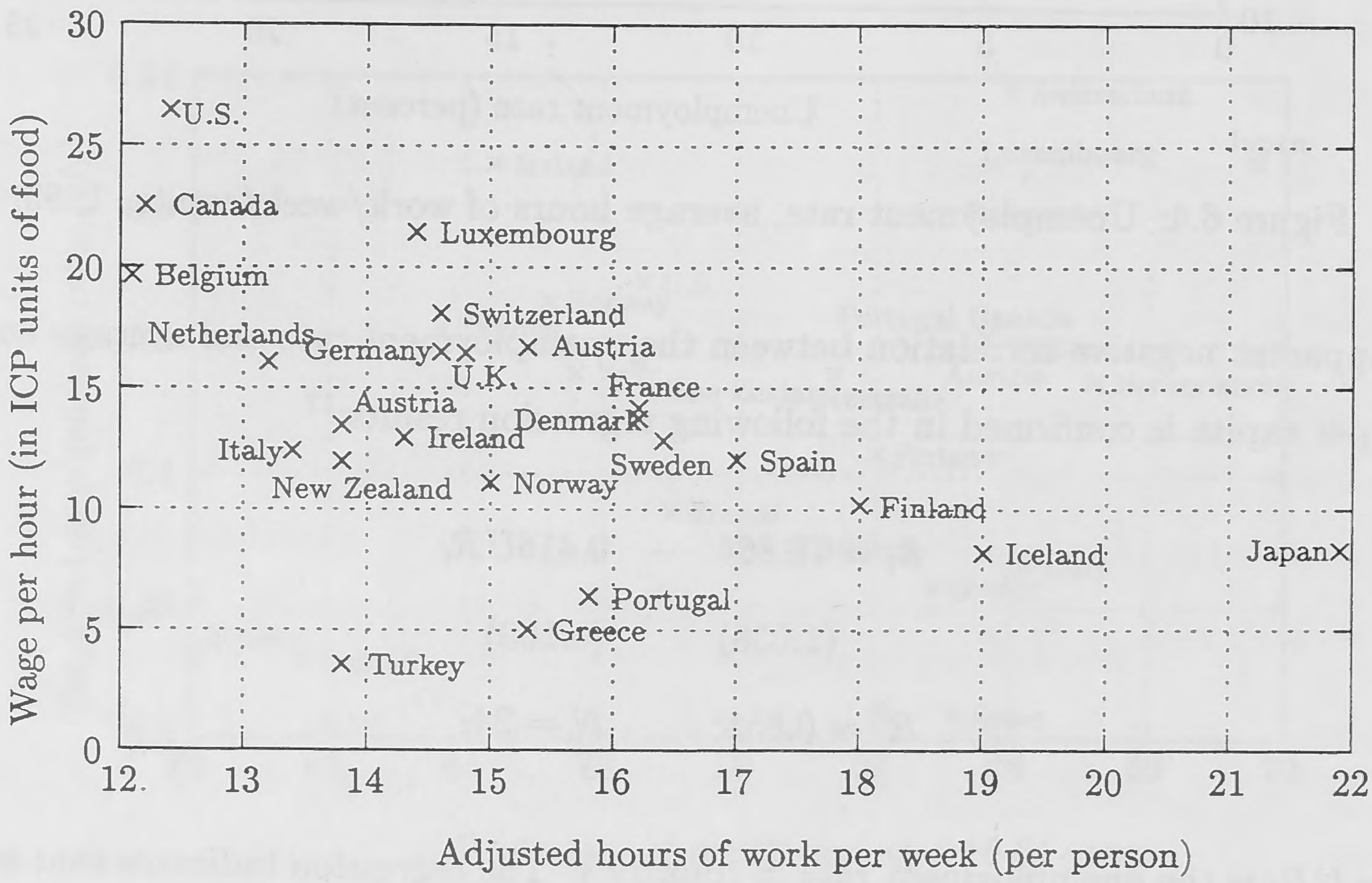

Figure 6.5: Hourly wages, adjusted hours of work, 1993

The bounds to subsistence leisure for the adjusted data were estimated using the same method described above. As shown in Figure 6.6, the estimated upper bound to γ_{K+1} was lowered when Belgium and France were added the AHARP set, while the lower bound was tightened significantly from 0 hours/year to 1936 hours/year when Spain was included. The final set of countries satisfying AHARP contained 18 countries; the maximum estimate of subsistence leisure calculated for this set was 2795 hours/year (or 7.7 hours/day).[13] This estimate of γ_{K+1} was used in the construction of the Approximate Leisure-Inclusive Ideal Afriat Index presented in the final two columns of Table 6.1.

[13]The minimum and mean estimates were 1937 and 2372 hours/year, respectively. It took 4170 iterations to collect 500 $\tilde{\gamma}$s in $\tilde{\mathcal{G}}$ for these 18 countries.

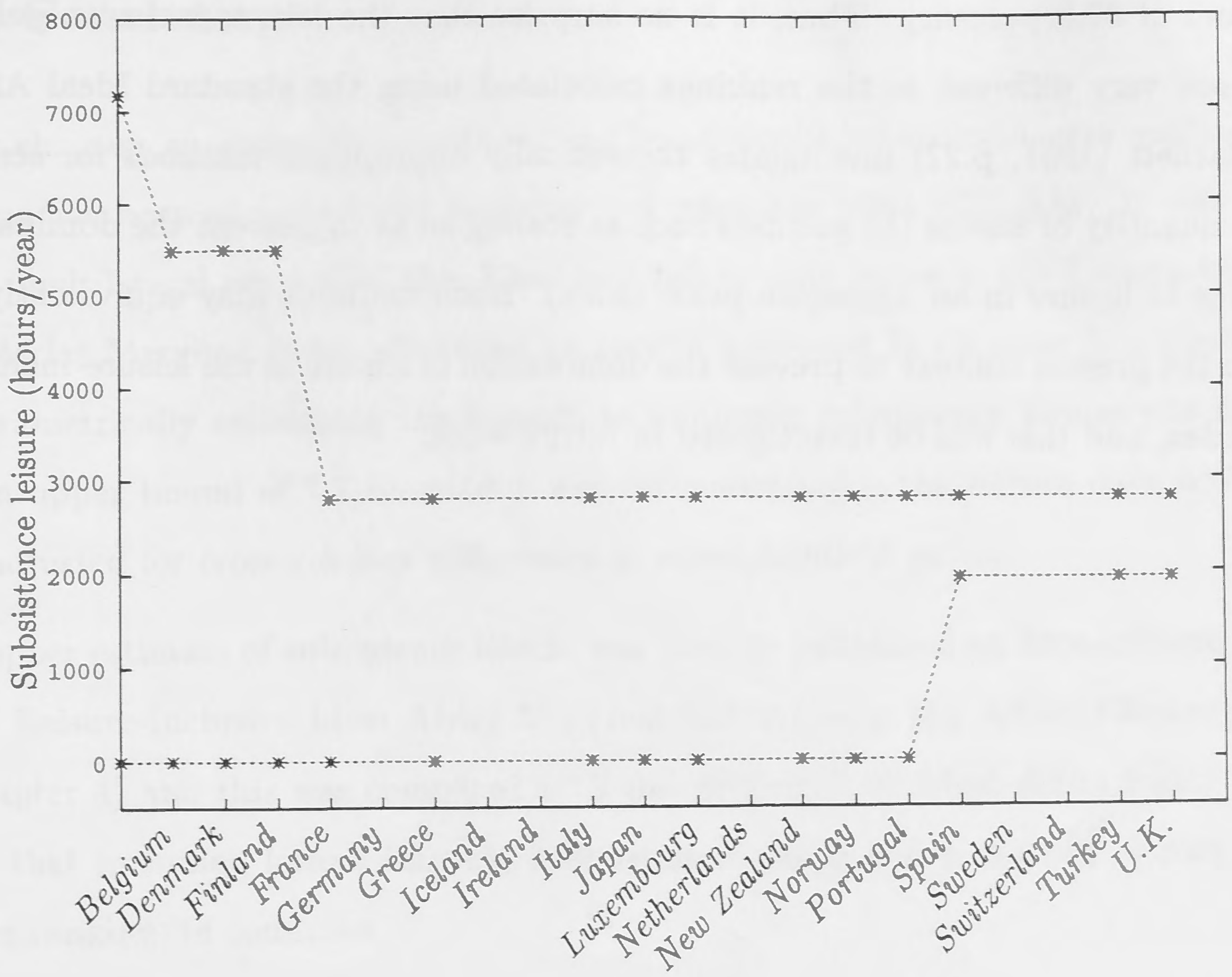

Figure 6.6: Estimated bounds to subsistence leisure - adjusted data, 1993

Leisure-inclusive welfare comparisons constructed using the adjusted leisure data are shown in the final two columns of Table 6.1. It is expected that $\tilde{\mathbf{a}}_m^*(0.99)$ calculated using the adjusted leisure data provides a more accurate leisure-inclusive welfare comparison across the OECD countries since the influence of differences in labour market conditions has been minimised. The rankings implied by $\tilde{\mathbf{a}}_m^*(0.99)$ calculated using the adjusted leisure data are closer to those based on $\mathbf{a}^*(0.98)$ (the Pearson's correlation coefficient of 0.823), but it is apparent that the inclusion of leisure still has a marked impact on how countries are ranked. Comparing the rankings implied by $\mathbf{a}^*(0.98)$ with those for $\tilde{\mathbf{a}}_m^*(0.99)$ calculated with the adjusted data, Japan slips from 9th to 19th and Iceland from 7th to 16th while Belgium moves up the rankings from 11th to 5th, and Italy moves up from 12th to 7th.

The reason that the inclusion of leisure has such a marked impact on cross-country welfare comparisons is that leisure accounts for a large share of full income. For the 24 OECD countries in 1993, the average share of supernumerary leisure in full income (calculated using the adjusted leisure data) was 82.7 percent (with a minimum of 71.7 percent and a maximum of 87.8 percent). Thus, it is no surprise that the leisure-inclusive welfare rankings are very different to the rankings calculated using the standard Ideal Afriat Index. Barnett (1981, p.27) investigates theoretically appropriate methods for scaling down the quantity of leisure (he justified such scaling so as to prevent the domination of the price of leisure in an aggregate price index). Such methods may equivalently be applied in the present context to prevent the domination of leisure in the leisure-inclusive welfare index, and this will be investigated in future work.

6.5 Conclusions

In this chapter, an approach for conducting leisure-inclusive cross-country welfare comparisons has been proposed and implemented using the 1993 ICP data. It was shown that a multilateral version of the Allen real full income index is the Leisure-Inclusive Ideal Afriat Marginal Index. Drawing on results presented in Chapter 5, a method for nonparametrically estimating the bounds to minimum subsistence leisure was applied, and an upper bound of 7.7 hours/day was estimated using the leisure data which had been adjusted for cross-country differences in unemployment rates.

This upper estimate of subsistence leisure was used to calculated an approximate version of the Leisure-Inclusive Ideal Afriat Marginal Index (using the Afriat efficiency index of Chapter 4) and this was compared with the Approximate Ideal Afriat Index. It was found that including leisure into the welfare comparison had a marked impact on the welfare rankings of countries.

There are three issues that need to be addressed in future research. First, because an indirect measure of leisure consumption is employed, differences in labour market conditions across countries are reflected in the leisure variable used in the analysis (countries with higher unemployment rates will record higher consumption of leisure). While an adjustment was made in attempt to purge the leisure data of the effects of cross-country differences in unemployment rates, this was essentially an ad hoc procedure. It would be preferable to instead use a shadow price of leisure (as used by Barnett (1981), although it may not be possible to construct such a price in a nonparametric setting. Alternatively, it may be possible to adapt the nonparametric test for rationing of Varian (1983, p.108) (and implemented by Swofford and Whitney (1994)) to the construction of true indexes in the presence of corner solutions. Second, it was apparent that particular countries can have a large impact on the estimate of subsistence leisure. An approach to methodically identify such outlier countries should be developed. Finally, because leisure accounts for a large share of full income, the leisure-inclusive welfare measures tend to be dominated by cross-country differences in leisure time consumed. An adaptation of the method for scaling down the quantity of leisure suggested by Barnett (1981) should be investigated.

6.5 Conclusions

In this chapter an approach for conducting leisure-inclusive cross-country welfare comparisons has been proposed and implemented using the 1993 ICP data. It was shown that a multilateral version of the Allen real full income index is the Leisure-Inclusive Ideal Afriat Marginal Index. Drawing on results presented in Chapter 5, a method for nonparametrically estimating the bound to minimum subsistence leisure was applied, and an upper bound of 7.7 hours/day was estimated using the leisure data which had been adjusted for cross-country differences in unemployment rates.

This upper estimate of subsistence leisure was used to calculated an approximate version of the Leisure-Inclusive Ideal Afriat Marginal Index (using the Afriat efficiency index of Chapter 4) and this was compared with the Approximate Ideal Afriat Index. It was found that including leisure into the welfare comparison had a marked impact on the welfare rankings of countries.

There are three issues that need to be addressed in future research. First, because an indirect measure of leisure consumption is employed, differences in labour market conditions across countries are reflected in the leisure variable used in the analysis (countries with higher unemployment rates will record higher consumption of leisure). While an adjustment was made in attempt to purge the leisure data of the effects of cross-country differences in unemployment rates, this was essentially an ad hoc procedure. It would be preferable to instead use a shadow price of leisure (as used by Barnett (1981)). Although it may not be possible to construct such a price in a nonparametric setting. Alternatively, it may be possible to adapt the nonparametric test for rationing of Varian (1983, p105) (and implemented by Swofford and Whitney (1994)) so the construction of true indexes in the presence of corner solutions. Second, it was apparent that particular countries can have a large impact on the estimate of subsistence leisure. An approach to methodically identify such outlier countries should be developed. Finally, because leisure accounts for a large share of full income, the leisure inclusive welfare measures tend to be dominated by cross-country differences in leisure time consumed. An adaptation of the method for scaling down the quantity of leisure suggested by Barnett (1981) should be investigated.

Conclusions

There were two main aims of this thesis. The first was to present a thorough review of the methods for conducting cross-country comparisons of welfare, with particular focus on the construction of multilateral true indexes. The second aim was to make three extensions to the use of multilateral true indexes in cross-country welfare comparisons. These extensions are now briefly summarised, and further research areas are identified.

Approximate multilateral true welfare comparisons

The test of HARP does not allow for the existence of measurement or consumer optimisation which may lead to a data set failing the test of common homothetic preferences. In Chapter 4, the work of Afriat(1972, 1987) and Varian (1993) was extended to the construction of Ideal Afriat Indexes which incorporate specific levels of consumer optimisation error.

A grid search method was used for finding the Afriat efficiency index e^* (which measures the degree of consumer optimisation error) necessary to include all countries in the welfare comparison. For the 1980 ICP data, consumer optimisation error of less than 5 percent allowed (approximate) welfare comparisons for 59 of the 60 countries, while for the 1993 ICP data 2 percent consumer optimisation error enabled all 24 countries to be included in the comparison. The use of approximate welfare comparisons allowed the identification of several pairwise rankings which reverse under differing levels of consumer optimisation error, in such cases it was suggested that the countries concerned should perhaps be ranked as equivalent.

The grid search method used for finding e^* is not particularly accurate, and future work will involve employing the iterative approach suggested by Houtman and Maks (1987) for estimating this parameter.

Multilateral true marginal welfare indexes

The Ideal Afriat Index exists when preferences are homothetic. However, homotheticity is a restrictive assumption and may not be appropriate in a cross-country context (where incomes typically vary greatly). In Chapter 5, it was shown that the Ideal Afriat Index is a special case of the Ideal Afriat Marginal Index, which exists when preferences are affine homothetic. The construction of the Ideal Afriat Marginal Index requires knowledge of the minimum subsistence consumption bundle; an iterative method for finding the bounds the set of possible subsistence bundles $\mathcal{G}$ was proposed. There is no empirical reason for basing welfare comparisons on a particular bundle in $\mathcal{G}$. Consequently, for a data set that satisfies HARP, there is no reason to use the Ideal Afriat Index for welfare comparisons rather than any other Ideal Afriat Marginal Index which can be constructed from another bundle in $\mathcal{G}$.

The overall bounds to the Ideal Afriat Marginal Indexes were constructed for the 1980 and 1993 data, and several inconsistencies between these bounds and the Ideal Afriat Index were identified. It was also shown that a comparison of the Ideal Afriat Index with the bounds to the Ideal Afriat Marginal Indexes can be used to classify countries as poor and rich.

The method used for finding bundles in $\mathcal{G}$ involves randomly selecting bundles from the uniform distribution (the bounds to which are appropriately set). However, a more satisfactory method would involve a grid search procedure and this will be investigated in future work.

Leisure-inclusive welfare comparisons

The aim of Chapter 6 was to propose a utility-consistent method for incorporating leisure into cross country comparisons of welfare. Results from Chapter 5 were used to show that a multilateral version of the Allen real full income index is in fact the Leisure-Inclusive Ideal Afriat Marginal Index. The iterative procedure suggested in Chapter 5 for finding the bounds to $\mathcal{G}$ was used to nonparametrically estimate minimum subsistence leisure (time spent eating, sleeping and performing other necessary biological functions). For the 1993 ICP data, the upper bound estimate of subsistence leisure (constructed using leisure data adjusted for cross-country differences in unemployment rates) was 7.7 hours/day.

This estimate was used in the construction of the Leisure-Inclusive Ideal Afriat Index and it was found that the welfare rankings implied by this index were significantly different to those found with the leisure-exclusive Ideal Afriat Index.

One of the problems with the approach to leisure-inclusive measures of welfare used in Chapter 6 is that cross-country differences in labour market conditions will affect the welfare comparison. In particular, a country with high unemployment will record a high average consumption of leisure, all other things equal, even though a lot of this will be involuntary leisure. An adjustment to the leisure variable to "purge" the effect of cross-country differences in unemployment rates was used, however it would be more theoretically acceptable to use a shadow price of leisure. The approach for calculating the shadow price of leisure suggested by Barnett (1981) may not be appropriate here as it involves a parametric specification, however it may be possible to adapt the nonparametric test for rationing suggested by Varian (1983) to the construction of true indexes in the presence of corner solutions. Another area for further research relates to the dominance of leisure in the Leisure-Inclusive Ideal Afriat Marginal Index; the method for scaling down the quantity of leisure suggested by Barnett (1981) should be investigated in the context of leisure-inclusive welfare measures.

This estimate was used in the construction of the Leisure-Inclusive Ideal Afriat Index and it was found that the welfare rankings implied by this index were significantly different to those found with the leisure-exclusive Ideal Afriat Index.

One of the problems with the approach to leisure-inclusive measures used in Chapter 6 is that cross-country differences in labour market conditions will affect the welfare comparison. In particular, a country with high unemployment will record a high average consumption of leisure, all other things equal, even though a lot of this will be involuntary leisure. An adjustment to the leisure variable to "purge" the effect of cross-country differences in unemployment rates was used, however it would be more theoretically acceptable to use a shadow price of leisure. The approach for calculating the shadow price of leisure suggested by Barnett (1981) may not be appropriate here as it involves a parametric specification, however it may be possible to adapt the nonparametric test for rationing suggested by Varian (1983) to the construction of true indexes in the presence of corner solutions. Another area for further research relates to the dominance of leisure in the Leisure-Inclusive Ideal Afriat Marginal Index; the method for scaling down the quantity of leisure suggested by Barnett (1981) should be investigated in the context of leisure-inclusive welfare measures.

Appendix A

Data Definitions and Sources

A.1 Prices and Quantities for Goods and Services

In this section, information on the source of the price and quantity data used in the thesis is provided.

A.1.1 The International Comparison Programme - background

The International Comparison Programme (ICP) is the source of the price and quantity data for goods and services.[1] The ICP provides detailed information on prices and goods which are comparable across countries. The first project director of the ICP was Professor Irving B. Kravis of the University of Pennsylvania and the first three phases (1970, 1973 and 1975) were coordinated from that university (see Table A.1).[2] After phase III the role of the University of Pennsylvania was gradually changed from main organiser of the ICP into that of adviser on methodological issues and developers of the Penn World Tables.

The first three phases were designed for global comparisons in that while some results were presented by region, the valuation of each country's quantities was carried out at "average" prices (to be defined below) of all participating countries. From 1980, however, the data has been organised by regional coordinators, and the regional results are linked together at a later stage for use in global comparisons. The global results for

[1]See World Bank (1993) and UN (1992) for a summary of the history and methods of the ICP; the following summary is based on these publications. Note that the letter "P" in the acronym ICP originally stood for Project, but this was changed at the twenty-fifth session of the Statistical Commission in 1989.

[2]The ICP is also associated with the Penn World Tables of Professors Alan Heston and Robert Summers.

Table A.1: Phases of the ICP and participation of countries

Phases (I-VII)	Developing countries	Industrialised countries	Total
1970	4	6	10
1973	8	8	16
1975	21	13	34
1980	42	18	60
1985	42	22	64
1990	6	24	30
1993	0	24	24

Note: Data for all countries participating in the 1993 phase of the ICP had not been released at the time of the completion of this thesis. Therefore these numbers are preliminary.

1980 were processed by UNSTAT and for 1985 by the Statistical Office of the European Communities (EUROSTAT), which has had an increasing role in the ICP.

A.1.2 A guide to using the ICP data

It is now described how the ICP data are constructed and how they may be used in cross-country comparative research.

The ICP implementation of the Geary method of aggregation

Different aggregation methods are used at different stages in the construction of the ICP data. However, the method used by the ICP to compute purchasing-power-parity-corrected exchange rates and real income measures is the Geary method, described in Chapter 1. The Geary method of aggregation was first suggested by Geary (1958), and is constructed under the assumed existence of "world prices" $\pi = (\pi_1, \ldots, \pi_K)$ and "true" exchange rates $\epsilon = (\epsilon^1, \ldots, \epsilon^N)$.[3]

The true exchange rates are Laspeyres price indexes, which compare the world prices with the prices of each country in turn:[4]

$$\epsilon^i = \frac{\sum_l \pi_l q_l^i}{\sum_l p_l^i q_l^i}, \quad i \in N, l \in K.$$

It is therefore apparent that each country's real income is the same whether valued at world prices $\sum_l \pi_l q_l^i$ or valued at domestic prices, converted at the true exchange rates

[3] The presentation here follows Neary (2000). The properties of the Geary index were investigated by Khamis (1972) and the Geary index is also known as the Geary-Khamis index.

[4] Following the U.S. convention, the ICP defines the true exchange rates or "purchasing power of currency" as the inverse of the definition presented above; the presentation in Neary (2000) follows the U.K. convention since it facilitates the matrix algebra.

$(\epsilon^i \sum_l p_l^i q_l^i)$. The world prices are defined by the requirement that total world spending on commodity l is the same whether valued at its world price $(\pi_l \sum_i q_l^i)$ or at domestic prices converted at the true exchange rates $(\sum_i \epsilon^i p_l^i q_l^i)$:

$$\pi_l = \frac{\sum_i \epsilon^i p_l^i q_l^i}{\sum_i q_l^i}, \quad i \in N, l \in K.$$

Solving simultaneously for $\boldsymbol{\epsilon}$ and $\boldsymbol{\pi}$, one can then calculate the real income of each country at world prices:[5]

$$x_G^i = \epsilon^i x^i = \sum_l \pi_l q_l^i, \quad i \in N, l \in K.$$

The ICP data are constructed from highly disaggregated data provided by the statistical agencies of the countries covered in each survey. These data are aggregated (using either the EKS method or the Cross Product Dummy (CPD) method) to *basic heading level* data. The requirements for basic heading level data are (a) expenditure in national currencies can be estimated for each basic heading (and are not available for the component items of the basic heading); and (b) the basic headings be as homogenous as possible in terms of dispersion of price ratios across countries.[6] The number of basic heading items across countries ranges between 150 and 258, depending on the amount of expenditure detail that is available for a particular country.

The basic heading level data (and aggregates constructed from these) are what are released to researchers and are what are used in the Geary aggregation equations above. In Table A.2 the expenditure aggregates constructed from the basic heading level data are presented; depending on the country, the basic heading levels lie either 1 or 2 levels below the lowest level of aggregation shown in Table A.2.

The Geary method was originally implemented by Geary using actual quantities of agricultural output and prices associated with these quantities. Thus, the basic heading level data for rice for India, for example, would be quantities measured in tons and prices measured in rupees per ton. The international prices computed by the method would be expressed in the numeraire currency (such as the U.S. dollar) and thus would be expressed as the number of dollars per unit quantity, for example, ton of rice.

[5]The solution can either be found via iteration or matrix algebra.

[6]For example, the dispersion of prices (relative to the U.S., for example) across countries for recreation equipment and services (category 1.7.1 in Table A.2) would be higher than the dispersion of prices of televisions, and thus televisions may be chosen as a basic heading level item.

Table A.2: ICP breakdown of GDP into expenditure aggregates

1	**Final household consumption**	1.7	*Recreation, entertainment, education, etc.*
1.1	*Food, beverages & tobacco*	1.7.1	Equipment and services
1.1.1	Food	1.7.2	Total education expenditures
1.1.2	Beverages	1.8	*Miscellaneous goods & services*
1.1.3	Tobacco	1.8.1	Personal care
1.2	*Clothing and footwear*	1.8.2	Other
1.2.1	Clothing	1.9	*Net expenditures of residents abroad*
1.2.2	Footwear	2	**Capital formation**
1.3	*Gross rent, fuel & power*	2.1	*Domestic capital formation*
1.3.1	Gross rent	2.1.1	Gross fixed capital formation
1.3.2	Fuel and power	2.1.2	Changes in stocks
1.4	*House furnishings, operations*	2.2	*Net foreign balance*
1.4.1	Furniture & appliances	3	**Government consumption, total**
1.4.2	Household goods and services	3.1	*Compensation of employees*
1.5	*Total medical care & services*	3.2	*Commodities, goods & services*
1.5.1	Private medical care & services	4	**Gross domestic product**
1.5.2	Public medical care		
1.6	*Transport and communication*		
1.6.1	Personal transportation equip.		
1.6.3	Purchased transport		
1.6.4	Communication		

However, in the ICP implementation of the Geary method, actual prices and quantities are not used in the above formulae. Rather, the basic heading level data of the ICP consist of expenditures x_l^i and *basic heading parities*, pp_l^i that have been computed either by the EKS or CPD methods. The basic heading parities are constructed as:

$$pp_l^i = p_l^i / p_l^{USA},$$

where the U.S. has been chosen as the base or numeraire country. This is the first of two normalisations made in the construction of the ICP data and the basic heading parities thus have the interpretation of units of currency of country i to the U.S. dollar.[7] It follows that the ICP quantities at the basic-heading level are not measured in physical units, but are termed *notional quantities* and are computed from the expenditures and basic heading parities as:

$$q_l^i = x_l^i / pp_l^i.$$

Each country's expenditure at the basic heading level is therefore converted into the currency of the numeraire country, and is thus termed a notional quantity.

[7] Note that this normalisation is only mentioned here for completeness since it doesn't affect the use of the ICP data. Some ICP publications (for example, World Bank (1993)), do not even mention that this normalisation has been made.

In the light of this normalisation, the ICP implementation of the Geary system can be re-written as (where q_l^i now refers to the notional quantity, not actual quantity):

$$\begin{aligned} \epsilon^i &= \frac{\sum_l \pi_l q_l^i}{\sum_l pp_l^i q_l^i}, \\ \pi_l &= \frac{\sum_i \epsilon^i pp_l^i q_l^i}{\sum_i q_l^i}, \quad i \in N, l \in K. \end{aligned} \tag{A.1}$$

Solving the Geary system

The Geary system can be solved either by use of matrix algebra or by iteration; the following is a description of the matrix algebra approach.

The system (A.1) consists of $N+K$ equations and $N+K$ unknown variables. The second normalisation in the Geary approach is to set the ϵ for the base country equal to one, and thus the exchange rate equation for the base country is dropped from the system. Assuming that the numeraire country is country N, the system (A.1) can be re-written as the following matrix system:

$$\begin{bmatrix} \hat{\mathbf{x}}_{-N} & -\mathbf{Q}'_{-N} \\ \mathbf{X}_{-N} & -\hat{\mathbf{q}} \end{bmatrix} \begin{bmatrix} \epsilon_{-N} \\ \pi \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\mathbf{x}^N \end{bmatrix},$$

where $\mathbf{x}$ denotes an $N \times 1$ vector of total expenditures by country, with typical element $x^i = \sum_l pp_l^i q_l^i$; $\mathbf{X}$ denotes the $K \times N$ matrix of expenditures by commodity and country, with typical element $X_{li} = x_l^i = pp_l^i q_l^i$; $\mathbf{x}^N$ is the $K \times 1$ vector of expenditures by commodity for country N (i.e., the final column of $\mathbf{X}$); $\mathbf{Q}$ denotes the $K \times N$ matrix of notional quantities by commodity and country, with typical element $Q_{li} = q_l^i$; and $\mathbf{q}$ denotes the $K \times 1$ vector of world consumption levels of each commodity, with typical element $q_l = \sum_i q_l^i$. Finally, a prime ($'$) denotes the transpose; a circumflex ($\hat{}$) over a vector denotes a diagonal matrix formed by placing on the principal diagonal the corresponding elements of the vector; and the subscript ($_{-N}$) denotes a vector or matrix from which the entries corresponding to the numeraire country (country N) have been deleted.

Noting that $\mathbf{W} \equiv \mathbf{Z}\hat{\mathbf{z}}^{-1}$ is the matrix of world budget shares (in domestic prices), with typical element $W_{li} = \omega_l^i = x_l^i / x^i$, then the above system can be solved for π using the formula for inverting a partitioned matrix:[8]

$$\boldsymbol{\pi} = [\hat{\mathbf{q}} - \mathbf{W}_{-N}\mathbf{Q}'_{-N}]^{-1}\mathbf{x}^N.$$

Interpreting the output of the Geary method

The normalisation $\epsilon^{US} = 1$ leads a particular interpretation of the international prices $\boldsymbol{\pi}$. An implication of this normalisation is that $x^{US} = \sum_l \pi_l q_l^{US}$. Therefore x^{US} international dollars can purchase U.S. GDP and hence 1 international dollar can purchase a quantity of goods and services equal to what 1 USD can purchase of U.S. GDP. Therefore, the international prices $\boldsymbol{\pi}$ are measured in international dollars (I\$) which have the same purchasing power as the U.S. dollar. However, note that while the international prices are denominated in U.S. dollars, by construction they reflect an "average international" price structure rather than U.S. price relativities.

The normalisation $\epsilon^{US} = 1$ also leads to a particular interpretation of the true exchange rates ϵ. The true exchange rate ϵ^i is defined as the number of units of country i's currency required to purchase the same amount of goods and services as 1 USD would buy in the U.S. This can be seen from the following example. Suppose $\epsilon^{AUS} = 1.5$ and $\sum_l \pi_l q_l^{AUS}$=15 I\$. Then from $\epsilon^{AUS} x^{AUS} = \sum_l \pi_l q_l^{AUS}$, it is apparent that 10 Australian dollars will purchase the same amount of goods and services as 15 I\$ and 1.5 Australian dollars will purchase the same amount of goods and services as 1 I\$. But because the I\$ has been normalised so 1 I\$ can purchase the same quantity of goods and services that 1 USD can purchase of U.S. GDP, then $\epsilon^{AUS} = 1.5$ is the number of Australian dollars required to purchase the same amount of goods and services as 1 USD would purchase in the U.S.

The Geary method of aggregation - an example

The following simple example is used to illustrate the Geary method and the use in the ICP of notional quantities and basic heading parities. Assume the ICP has collected basic heading data on four commodities (bread, cheese, footwear and jumpers) for three countries (Australia, India and the U.S.). The basic heading raw data consists of expenditures in local currencies on the different commodities (the sum of which comprise

[8] See, for example, Greene (1993, p.27).

Table A.3: Example - basic heading level expenditure, local prices and budget shares

	x_l^i			p_l^i			ω_l^i		
	AUS	IND	USA	AUS	IND	USA	AUS	IND	USA
Food	*7000*	*36000*	*6400*						
Bread	3500	24000	1600	1.50	20.00	2.00	0.25	0.40	0.10
Cheese	3500	12000	4800	2.00	60.00	1.00	0.25	0.20	0.30
Clothing	*7000*	*24000*	*9600*						
Footwear	4200	21000	3200	8.00	100.00	6.00	0.30	0.35	0.20
Jumpers	2800	3000	6400	12.00	150.00	10.00	0.20	0.05	0.40
GDP	**14000**	**60000**	**16000**				1.00	1.00	1.00

Table A.4: Example - basic heading parities and notional quantities

	pp_l^i			q_l^i		
	AUS	IND	USA	AUS	IND	USA
Food						
Bread	0.75	10.00	1.00	4667	2400	1600
Cheese	2.00	60.00	1.00	1750	200	4800
Clothing						
Footwear	1.33	16.67	1.00	3150	1260	3200
Jumpers	1.20	15.00	1.00	2333	200	6400
GDP				**11900**	**4060**	**16000**

GDP) and the local prices per unit (Table A.3)..

The basic heading data used in the Geary aggregation are the parities (where the U.S. is chosen as the numeraire country) and notional quantities, defined above (and shown in Table A.4). A question one may ask is why can't we just use the sum of the notional quantities as the real income measure for each country? While this would express GDP in a common currency, the problem is that real GDP would be being measured at U.S. prices and it is preferable that real income does not depend on the prices of any one country.

The prices and quantities in the above table are used as input into the Geary method, and the following international prices and true exchange rates are calculated:

$$\pi = [0.629, 1.188, 0.973, 0.965], \qquad \epsilon = [0.738, 0.053, 1.000].$$

The Geary estimate of real income for country i can then be calculated as $\epsilon^i x^i$ or $\sum_l \pi_l q_l^i$ (Table A.5). Table A.5 also presents the pp_l^i/π_l, which are called the *purchasing power parities* for each commodity. One of the main advantages of the Geary method is that it exhibits matrix consistency in that the real income measure can be

Table A.5: Real income using the Geary method

	pp_l^i/π_l			$\pi_l q_l^i$		
	AUS	IND	USA	AUS	IND	USA
Food	*1.396*	*20.612*	*0.954*	*5013*	*1747*	*6706*
Bread	1.193	15.904	1.590	2934	1509	1006
Cheese	1.684	50.523	0.842	2078	238	5700
Clothing	*1.316*	*16.908*	*1.033*	*5319*	*1419*	*9294*
Footwear	1.370	17.123	1.027	3066	1226	3115
Jumpers	1.243	15.537	1.036	2253	193	6179
GDP	0.738	0.053	1.000	**10331**	**3166**	**16000**

consistently disaggregated by commodity. For example, in Table A.5, the real value of food consumption in Australia is calculated as the sum of the real value of consumption of bread and cheese. Thus, while the Geary method does not derive an international price of food (since food is not a basic heading item), the fact the method exhibits matrix consistency enables one to implicitly calculate $q_{food}^{AUS}\pi_{food}$. Similarly, the purchasing power parity of food for Australia, $pp_{food}^{AUS}/\pi_{food}$ can be calculated as $pp_{food}^{AUS}/\pi_{food} = x_{food}^{AUS}/q_{food}^{AUS}\pi_{food} = 7000/5013 = 1.396$. The fact that the Geary method is matrix consistent therefore allows one to derive consistent prices and quantities for aggregation levels higher than the basic heading level.

Note that the true exchange rates or purchasing power of currencies ϵ are also shown in Table A.5 (last row, columns 3-4). This reflects the fact that they can be calculated as (the inverse of) a weighted average of the purchasing power parities, using the budget shares as weights:

$$(\epsilon^i)^{-1} = \sum_l \frac{\pi_l}{pp_l^i}\omega_l^i.$$

The data provided by the ICP (i.e. in hardcopy and in the *STARS* software) are the real expenditures $q_l^i\pi_l$, purchasing power parities pp_l^i/π_l and purchasing power of currencies ϵ^i shown in Table A.5, as well as the nominal expenditures shown in Table A.3. The real expenditures and purchasing power parities are often referred to as *ICP quantities and prices*, and are used in international comparison work. For example, Dowrick and Quiggin (1997) use the basic heading level ICP quantities and prices in their work on constructing multilateral true indexes.

Note, however, that Dowrick and Quiggin (1997) could equivalently have used the basic heading parities pp_l^i and notional quantities q_l^i rather than the ICP prices and quantities.

Following Neary (2000), these can be recovered from the example ICP data above in the following way. At the basic heading level, dividing each purchasing power parity pp_l^i/π_l by the corresponding entry for the U.S. pp_l^{US}/π_l gives $pp_l^i/pp_l^{US} = pp_l^i$, since by definition $pp_l^{US} = 1$. The notional quantity is then recovered as $q_l^i = x_l^i/pp_l^i$.

Using this method, it is possible to compute parities and notional quantities for more aggregated expenditure levels. In the above example, the parities and notional quantities were provided only at the basic heading level. However, the parity for food in Australia can be calculated as $(pp_{food}^{AUS}/\pi_{food})/(pp_{food}^{USA}/\pi_{food}) = 1.396/0.954 = 1.463$. Similarly, the notional quantity of food in Australia is calculated as $q_{food}^{AUS} = x_{food}^{AUS}/pp_{food}^{AUS} = 7000/1.463 = 4785$.

The multilateral methods of Dowrick and Quiggin (1997) can be equivalently applied to the ICP prices and quantities or the parities and notional quantities. The reason for this is that the log Laspeyres matrix will be identical regardless of which set of data is used.[9]

A.1.3 1980 data

Following from the above discussion, it is now possible to describe the 1980 ICP data used in this thesis. The 1980 data on prices and quantities of goods and services are from Phase IV of the ICP, and are available in hard copy in World Bank (1993) and electronically via the World Bank's *STARS* software. The 1980 ICP data cover 60 countries and are available at a very disaggregated level.

Table 6 of World Bank (1993) gives data for 1980 on per capita expenditure on different types of goods and services measured in national currencies, $x_l^i = pp_l^i q_l^i$, where $pp_l^i = p_l^i/p_l^{USA}$ is the parity of good l in country i (with the U.S. as the base country) and $q_l^i = x_l^i/pp_l^i$ is the per capita notional quantity of good l consumed in country i. Table 4 of World Bank (1993) gives expenditures measured in international or world prices, $\pi_l q_l^i$, where π_l is the international price of good l (these prices are outputs of the Geary method of aggregation). The quantities in Table 4 are therefore measured in "international dollars".

Table 5 of World Bank (1993) gives purchasing power parities, which are the nominal

[9] Further, this property will hold regardless of whether basic heading level data or more aggregate data is being used.

expenditures in Table 6 divided by the real expenditures in Table 4, $(pp_l^i q_l^i)/(\pi_l q_l^i) = pp_l^i/\pi_l$. Thus, Table 4 gives the ICP quantities $(\pi_l q_l^i)$ and Table 5 gives the ICP prices (pp_l^i/π_l) which can be used to construct real income measures (these basic heading level data were used by Dowrick and Quiggin (1997), for example).

The analysis in the thesis is conducted at a fairly aggregated level (the level of aggregation covers 18 types of goods and services), and for reasons discussed in Chapter 1, only includes components of household expenditure.[10] It should be noted that because aggregated, rather than basic heading level data are used in this thesis, the data used are therefore an output of the Geary method of aggregation (and are derived using the method explained above). Thus, it is somewhat contradictory that while the methods of aggregation studied in this thesis are "competitors" to the Geary method, the latter method has been used in the preparation of the data which is used in the empirical implementation. However, aggregated data were only used for computational and presentational ease (and to avoid further issues with missing prices); all the methods discussed in this thesis could equally be applied to the basic heading level data. Further, while the Geary method is used to derive the consumption aggregates, they are still relatively disaggregated and thus it can be argued that the use of the Geary method is not significantly "tainting" the results.

Even at this relatively high level of aggregation, two countries (Bolivia and Sri Lanka) have missing price and quantity information. Bolivia is missing these data for the "other" category of expenditure and Sri Lanka is missing the data for the footwear category. So as to be able to keep these countries in the data set, it was decided to make an imputation for the missing data. It is possible that a country recorded zero expenditure on a particular good or service, and hence no imputation was made for the missing quantities. Two methods were used to impute for the missing prices. The missing price of footwear for Sri Lanka was imputed as the following:

$$\tilde{p}^{80}_{footwear} = \frac{p^{75}_{footwear}}{p^{75}_{clothing\&footwear}} \times p^{80}_{clothing\&footwear}.$$

Thus, the imputed price of footwear for 1980 reflects the same relativity between the price of footwear and the price at the next higher level of aggregation (footwear and clothing)

[10] Note that at this level of aggregation, the 1980 ICP data include quantities and prices for public medical expenditure, but since these data were missing for several countries it has not been included in the present study.

which was observed in 1975. Using this method, the imputed 1980 price of footwear for Sri Lanka is $\tilde{p}^{80}_{footwear} = 2.732/3.248 \times 5.069 = 4.264$.

Bolivia was not in the 1975 data set, and hence the same method could not be used to impute for the price of "other" goods. Consequently, the imputed price of "other" goods was set equal to the price at the next higher level of aggregation, which is miscellaneous goods and services.

Table A.6 presents the descriptive statistics of the ICP price and quantity data for the 60 countries.

A.1.4 1993 data

The 1993 data are from phase VII of the ICP and cover 24 OECD countries (at present the 1993 data can only be used for regional, rather than global comparisons). At the 19-good level of disaggregation, three countries (Belgium, Germany and the Netherlands) had missing prices for public medical care; these missing prices were imputed for using the approach outlined above. Table A.7 presents the descriptive statistics of the price and quantity data for the 24 OECD countries.

A.2 Leisure Time, Wages and Demographic Data

In Chapter 6, data on leisure time and wages are merged with the 1993 ICP data to compute leisure-inclusive welfare indexes (see Table A.8 for the descriptive statistics of these data).

The data for leisure and wages were constructed using the methods outlined in Dowrick and Quiggin (1994). The number of hours worked per week per capita (h^w) is calculated:

$$h^w = e/p \times \textit{average hours per worker per week},$$

where e/p is the employment-to-population ratio derived from Table 1 of OECD (1997). Average hours per worker per week is derived from Table 4A of the 1997 edition of the ILO's Yearbook of Labour Statistics. This variable measures average hours *worked* in all sectors of the economy, as opposed to average hours *paid for* (which includes annual leave and public holidays).[11]

[11]Some countries only recorded average hours paid; an imputation was made for these countries. Where available, average hours worked by employees was used, but for some countries (e.g. the U.S.) only average hours worked by wage earners was available.

Table A.6: 1980 ICP data ($N = 60, K = 18$)

	Quantities				*Prices*			
	mean	min.	max.	s				
1.1.1 Food	679.5	75.8	1505.0	394.6	1.000	1.000	1.000	0.000
1.1.2 Beverages	97.1	0.4	307.3	101.5	1.500	0.353	4.563	0.828
1.1.3 Tobacco	47.1	1.2	184.0	42.6	1.240	0.329	2.446	0.543
1.2.1 Clothing	178.6	10.8	855.0	170.8	1.068	0.440	2.374	0.375
1.2.2 Footwear	37.1	0.0	116.4	28.1	1.082	0.418	1.940	0.364
1.3.1 Gross rent	279.7	1.7	1072.1	294.9	1.370	0.309	4.248	0.812
1.3.2 Fuel & power	96.8	2.7	588.2	118.1	1.256	0.260	2.839	0.577
1.4.1 Furnishing & appliances	124.6	1.4	591.3	142.1	1.139	0.466	2.453	0.375
1.4.2 Household goods & services	87.6	4.4	245.0	62.2	0.994	0.153	2.090	0.443
1.5.1 Private medical care & services	190.2	3.1	758.8	197.0	1.001	0.234	2.250	0.414
1.6.1 Personal transport equipment	74.9	0.1	490.5	119.9	1.744	0.391	6.723	1.133
1.6.2 Operation of transport equipment	132.6	0.9	773.9	170.9	1.237	0.302	2.248	0.380
1.6.3 Purchased transport	50.0	0.7	260.3	54.9	1.422	0.279	2.897	0.685
1.6.4 Communication	29.8	0.1	181.9	36.9	1.461	0.251	4.482	0.881
1.7.1 Recreation equipment & services	157.4	0.9	675.3	169.1	1.120	0.478	3.634	0.444
1.7.2 Education expenditure	180.5	14.1	503.9	131.7	0.910	0.207	2.246	0.604
1.8.1 Personal care	90.6	0.5	722.3	111.9	1.096	0.565	2.054	0.337
1.8.2 Other miscellaneous goods	214.3	0.0	923.5	266.6	1.052	0.201	2.571	0.463

Note: In this table, prices are expressed in ICP units of food.

Table A.7: 1993 ICP data ($N = 24, K = 19$)

	Quantities				*Prices*			
	mean	min.	max.	s				
1.1.1 Food	1737.0	1081.8	2747.8	354.4	1.000	1.000	1.000	0.000
1.1.2 Beverages	309.2	13.4	480.6	100.2	1.102	0.784	1.587	0.267
1.1.3 Tobacco	233.9	84.4	1213.5	215.4	1.187	0.457	2.009	0.366
1.2.1 Clothing	596.3	271.8	1082.6	185.2	1.129	0.841	1.594	0.204
1.2.2 Footwear	127.5	53.3	260.9	53.2	1.126	0.774	2.163	0.284
1.3.1 Gross rent	1922.3	876.4	2710.7	556.6	0.971	0.362	1.598	0.266
1.3.2 Fuel & power	511.6	165.0	1422.7	306.2	0.989	0.353	1.429	0.284
1.4.1 Furnishing & appliances	458.5	193.4	996.1	186.0	1.114	0.768	1.576	0.189
1.4.2 Household goods & services	371.1	162.7	692.8	134.9	1.033	0.762	1.533	0.176
1.5.1 Private medical care & services	1087.0	165.4	2534.4	825.5	0.852	0.466	1.605	0.220
1.5.1 Public medical care & services	620.3	0.0	1669.5	488.9	0.973	0.550	1.926	0.260
1.6.1 Personal transport equipment	436.2	49.5	1912.1	378.6	1.321	0.579	2.592	0.445
1.6.2 Operation of transport equipment	641.9	51.8	1428.0	284.2	1.203	0.877	1.500	0.144
1.6.3 Purchased transport	310.7	113.9	791.5	151.8	0.964	0.442	1.360	0.251
1.6.4 Communication	181.8	26.4	403.1	93.5	1.238	0.328	2.846	0.512
1.7.1 Recreation equipment & services	865.8	79.9	1804.6	375.5	1.214	0.838	1.667	0.197
1.7.2 Education expenditure	1254.9	439.4	1932.9	347.6	0.842	0.432	1.407	0.236
1.8.1 Personal care	368.6	122.8	908.7	167.0	1.091	0.830	1.652	0.182
1.8.2 Other miscellaneous goods	1784.9	441.1	3191.9	652.5	1.017	0.755	1.300	0.136

Note: In this table, prices are expressed in ICP units of food.

Table A.8: Data for analysis in Chapter 6, 1993

	mean	min.	max.	s
e/p	0.43	0.31	0.55	0.06
Hours/worker/week	35.1	24.8	47.5	4.9
h^w	15.1	11.2	24.9	3.0
Wage share	0.521	0.312	0.642	0.068
GDP (in ICP units of food)	19811.3	5595.0	32876.0	5754.3
w (in ICP units of food per hour)	13.76	3.57	26.46	5.53
$p^{15\text{-}64}$	66.4	62.0	69.8	2.0
UR	9.0	2.5	22.7	4.6

Leisure time per year (l) is calculated:

$$l = (T \times 7 - h^w) \times 52,$$

where T is the total amount of time that is split between working and leisure.

The hourly wage rate (w) per capita is calculated:

$$w = \frac{GDP\ per\ capita \times wage\ share}{(h^w \times 52)},$$

where GDP per capita is 1993 per capita nominal GDP from the ICP and the wage share is derived from OECD (1996) as the ratio of compensation of employees paid by resident producers to GDP (in current prices).

The standardised unemployment rate (UR) is from Annex Table 22 of OECD (2000).[12] The percentage of the population aged 15-64 years ($p^{15\text{-}64}$) is from Table 1 of OECD (1997).

A.3 Data Sources

Organisation for Economic Co-operation and Development (1997), *Labour Force Statistics, 1976-96*, Paris.

Organisation for Economic Co-operation and Development (1996), *National Accounts, 1960-94, Main Aggregates (Vol. I)*, Paris.

Organisation for Economic Co-operation and Development (2000), *OECD Economic Outlook*, Paris.

United Nations (1986,7), *World Comparisons of Purchasing Power and Real Product for*

[12]Standardised unemployment rates were not available for three countries - the unemployment rate constructed using the country specific methodology (Annex Table 21) was used for these countries.

1980, Vols I and II, United Nations and Commission of the European Communities.

United Nations (1992), *Handbook of the International Comparison Programme*, ST/ESA/STAT/SER.F/62, New York.

The World Bank (1993), *Purchasing Power of Currencies: Comparing National Incomes Using ICP Data*, Washington, D.C.

Appendix B

Summary of Mathematical Results and Geometric Concepts

This Appendix contains a summary of mathematical results and geometric concepts referred to in the text.

B.1 Mathematical Results

The Kuhn-Tucker Theorem

Consider the problem:[1]

$$\max_{\mathbf{x}} F(\mathbf{x}) \text{ subject to } G(\mathbf{x}) \leq \mathbf{c}, \mathbf{x} \geq 0$$

Form the Lagrangian:

$$L(\mathbf{x}, \lambda) = F(\mathbf{x}) + \lambda[\mathbf{c} - G(\mathbf{x})]$$

Suppose $\mathbf{x}^*$ maximises $F(\mathbf{x})$ subject to $G(\mathbf{x}) \leq \mathbf{c}$ and $\mathbf{x} \geq 0$, and the constraint qualification holds. Then there is a value of λ such that:

$$L_{\mathbf{x}}(\mathbf{x}^*, \lambda) \leq 0, \;\; \mathbf{x} \geq 0, \text{ with complementary slackness}$$
$$L_{\lambda}(\mathbf{x}^*, \lambda) \geq 0, \;\; \lambda \geq 0, \text{ with complementary slackness}$$

If $\mathbf{x}$ is constrained to be positive, then the Khun-Tucker conditions become:

$$L_{\mathbf{x}}(\mathbf{x}^*, \lambda) = 0, \;\; \mathbf{x} > 0$$
$$L_{\lambda}(\mathbf{x}^*, \lambda) \geq 0, \;\; \lambda \geq 0, \text{ with complementary slackness}$$

[1] This presentation is from Dixit (1990).

If the $G(\mathbf{x})$ constraints are assumed to hold with equality, then the Khun-Tucker conditions become:

$$L_{\mathbf{x}}(\mathbf{x}^*, \lambda) \leq 0, \;\; \mathbf{x} \geq 0, \text{ with complementary slackness}$$
$$L_{\lambda}(\mathbf{x}^*, \lambda) = 0, \;\; \lambda > 0$$

If $\mathbf{x}$ is constrained to be positive and the $G(\mathbf{x})$ constraints hold with equality, then the Khun-Tucker conditions become:

$$L_{\mathbf{x}}(\mathbf{x}^*, \lambda) = 0, \;\; \mathbf{x} > 0$$
$$L_{\lambda}(\mathbf{x}^*, \lambda) = 0, \;\; \lambda > 0$$

The Envelope Theorem

Consider the problem:[2]

$$\max_{\mathbf{x}} F(\mathbf{x}, \mathbf{a}) \text{ subject to } G(\mathbf{x}, \mathbf{a}) \leq \mathbf{c}, \mathbf{x} \geq 0,$$

where $\mathbf{x}$ is a vector of choice variables and $\mathbf{a}$ is a vector of parameters which may enter the objective function, the constraint, or both. The solution will clearly depend upon $\mathbf{a}$; denote this solution $\mathbf{x}(\mathbf{a})$.

Define a new function $M(\mathbf{a})$, which is the value achieved by the objective function when $\mathbf{x}$ is chosen to maximise F subject to the constraints. $M(\mathbf{a})$ is the *maximum value function* and is defined:

$$M(\mathbf{a}) \equiv \max_{\mathbf{x}} F(\mathbf{x}, \mathbf{a}) \text{ subject to } G(\mathbf{x}, \mathbf{a}) \leq \mathbf{c}, \mathbf{x} \geq 0$$

The maximum value function can equivalently be defined as:

$$M(\mathbf{a}) \equiv F(\mathbf{x}(\mathbf{a}), \mathbf{a}),$$

where $\mathbf{x}(\mathbf{a})$ is the optimal solution to the problem. The Lagrangian for this problem is:

$$L(\mathbf{x}, \mathbf{a}, \lambda) = F(\mathbf{x}, \mathbf{a}) + \lambda[\mathbf{c} - G(\mathbf{x}, \mathbf{a})]$$

Let $(\mathbf{x}(\mathbf{a}), \lambda(\mathbf{a}))$ solve the Khun-Tucker conditions. The Envelope Theorem states:

$$\frac{\partial M(\mathbf{a})}{\partial a_i} = \left.\frac{\partial L}{\partial a_i}\right|_{\mathbf{x}(\mathbf{a}),\lambda(\mathbf{a})} = \left.\frac{\partial F(\mathbf{x}, \mathbf{a})}{\partial a_i}\right|_{\mathbf{x}(\mathbf{a})} - \lambda(\mathbf{a}) \left.\frac{\partial G(\mathbf{x}, \mathbf{a})}{\partial a_i}\right|_{\mathbf{x}(\mathbf{a}),\lambda(\mathbf{a})}$$

[2]This presentation is from Jehle (1991).

Roy's Identity

Let $\psi(x, \mathbf{p})$ be an indirect utility function. Then:

$$\mathbf{q}_i(x, \mathbf{p}) = -\frac{\partial \psi(x, \mathbf{p}) \partial p_i}{\partial \psi(x, \mathbf{p}) \partial x}$$

Shephard's Lemma

Let $e(U, \mathbf{p})$ be a differentiable expenditure function. Then:

$$\frac{\partial e(U, \mathbf{p})}{\partial p_i} \equiv h_i(U, \mathbf{p}) = \mathbf{q}_i(x, \mathbf{p})$$

Homogeneous Functions

A real valued function $f(\mathbf{x})$ is called *homogeneous of degree k* iff for all $\theta > 0$:

$$f(\theta \mathbf{x}) \equiv \theta^k f(\mathbf{x}).$$

It is worth noting two special cases. The function $f(\mathbf{x})$ is *homogeneous of degree 1* or *linearly homogeneous* iff for all $\theta > 0$:

$$f(\theta \mathbf{x}) \equiv \theta f(\mathbf{x}).$$

The function $f(\mathbf{x})$ is *homogeneous of degree zero* iff for all $\theta > 0$:

$$f(\theta \mathbf{x}) \equiv f(\mathbf{x}).$$

It can be shown (see, for example, Jehle (1991, p.71)) that if $f(\mathbf{x})$ is homogeneous of degree k, then its partial derivatives are homogeneous of degree $k - 1$:

$$\frac{\partial f(\theta \mathbf{x})}{\partial x_l} = \theta^{k-1} \frac{\partial f(\mathbf{x})}{\partial x_l}.$$

An important implication of this is that the slopes of the level surfaces of a homogeneous function are constant along rays through the origin:

$$\frac{\frac{\partial f(\theta \mathbf{x})}{\partial x_l}}{\frac{\partial f(\theta \mathbf{x})}{\partial x_m}} = \frac{\frac{\partial f(\mathbf{x})}{\partial x_l}}{\frac{\partial f(\mathbf{x})}{\partial x_m}},$$

for all $\theta > 0$. Note, however, that certain non-homogeneous functions will have this same property. For example, affine transformations of homogeneous functions are themselves not homogenous, yet the slopes of the level surfaces are constant along rays through the origin.

B.2 Geometric Concepts

Convex hulls

The set of all the limit points of a set S is called the *derived set* of S. The union of S and its derived set is called the *closure* of S.

For a given set S a *convex hull* is the smallest convex set containing S.

For a given finite set S, a *polytope* is the smallest set containing S. A polytope is convex if and only if any line containing an edge of the polytope does not contain any interior points (a convex polytope is also known as a convex polyhedron). A concave polytope is any polytope that is not convex (see Figures (B.1) and (B.2)).

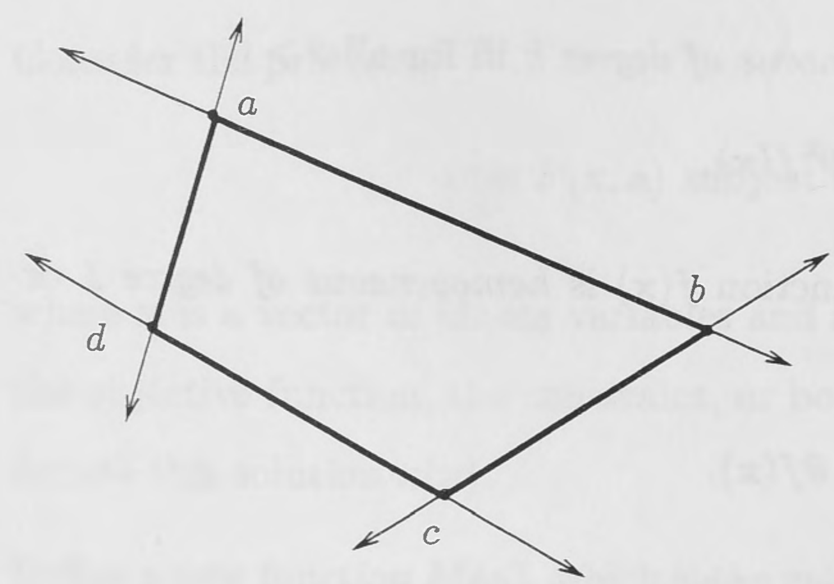

Figure B.1: A convex polytope

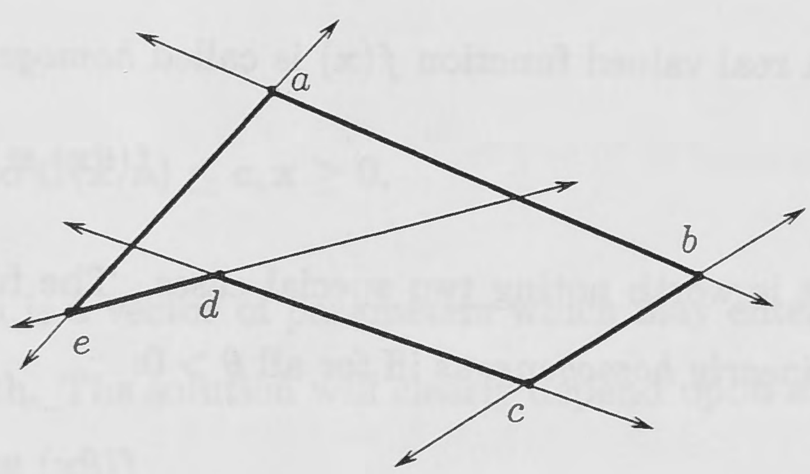

Figure B.2: A concave polytope

Cones

K is a *cone* if $\alpha \geq 0$ and $\mathbf{q} \in K$ implies $\alpha\mathbf{q} \in K$. Thus in $\mathbb{R}^2$, an example of K is a *half line* or ray starting from the origin. K is a *convex cone* if it is a cone with the additional property that $x, y \in K$ implies $x + y \in K$. Thus the set of two different half lines starting from the origin is a cone, but not a convex cone (the set $X \cup Y$ in Figure B.3 is an example of a cone, but not a convex cone). If the set includes the area inside two half lines with an acute angle, then it is a convex cone (Figure B.4).

For a given finite set S, a *polyhedral* is the smallest cone containing S while a *convex polyhedral* is the smallest convex cone containing S. Thus, in $\mathbb{R}^2$ a polyhedral consists of the smallest number of half lines starting from the origin which include all the points in S (in Figure B.5, the set $X \cup Y \cup Z$ is a polyhedral). A convex polyhedral includes the area

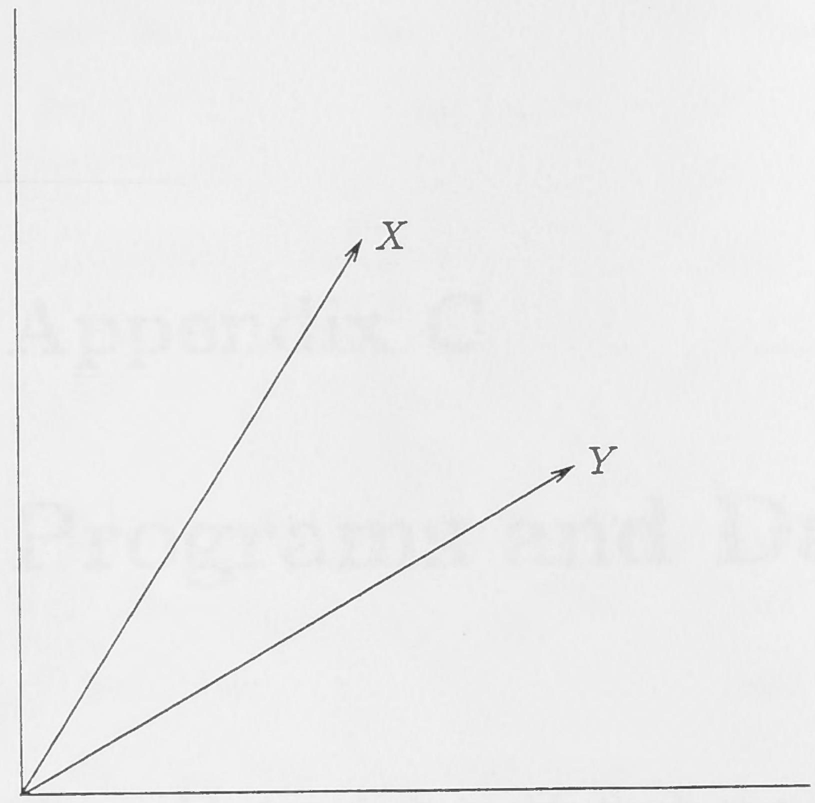

Figure B.3: A cone

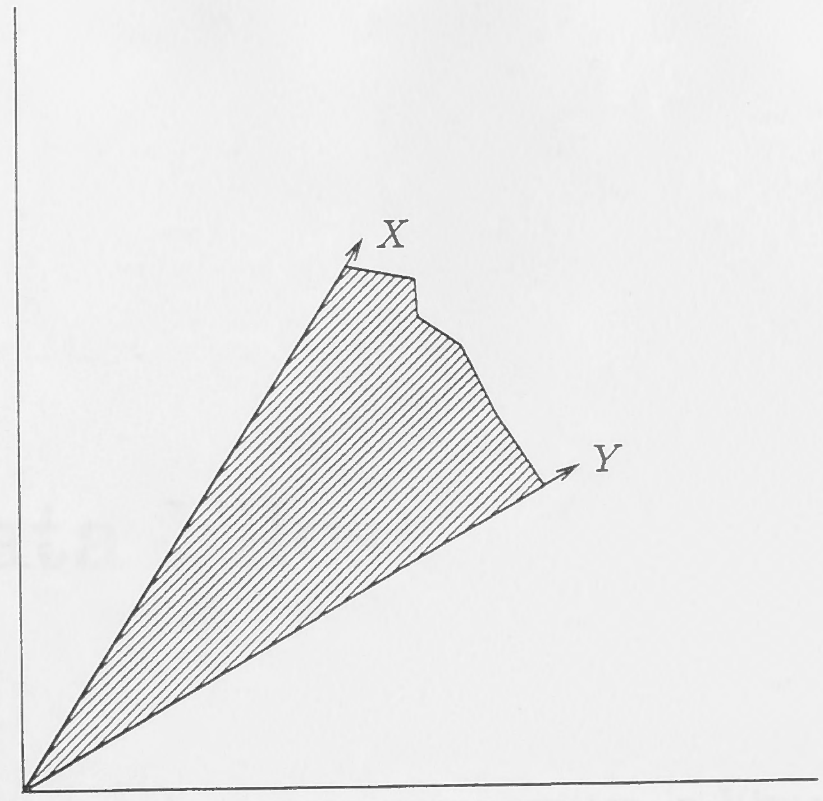

Figure B.4: A convex cone

inside the two half lines with the smallest acute angle which results in S being a subset of the convex polyhedral (Figure B.6). Finally, a concave polyhedral is any polyhedral which is not convex.

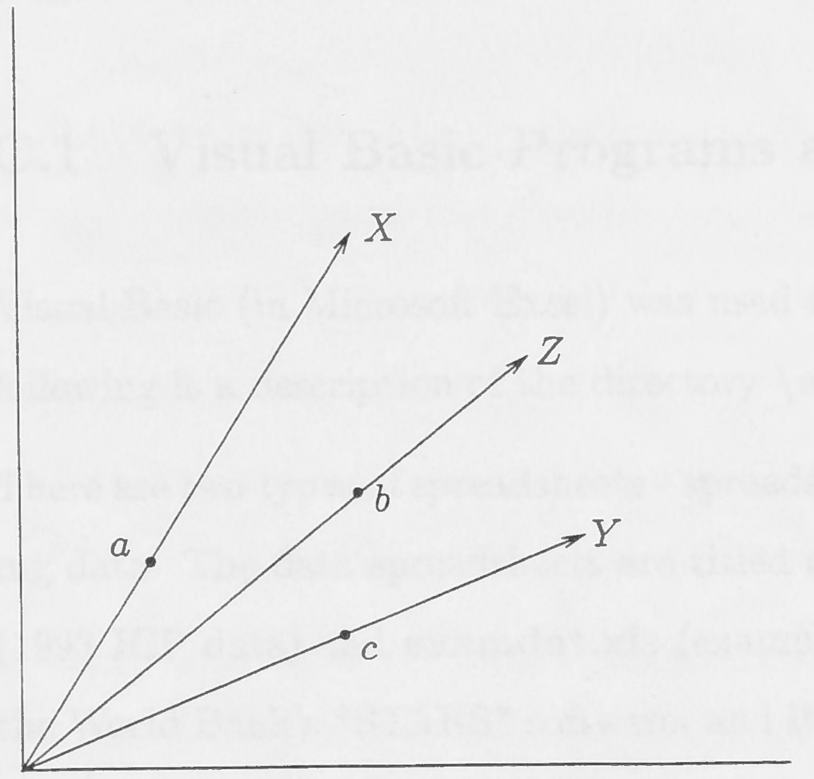

Figure B.5: A polyhedral

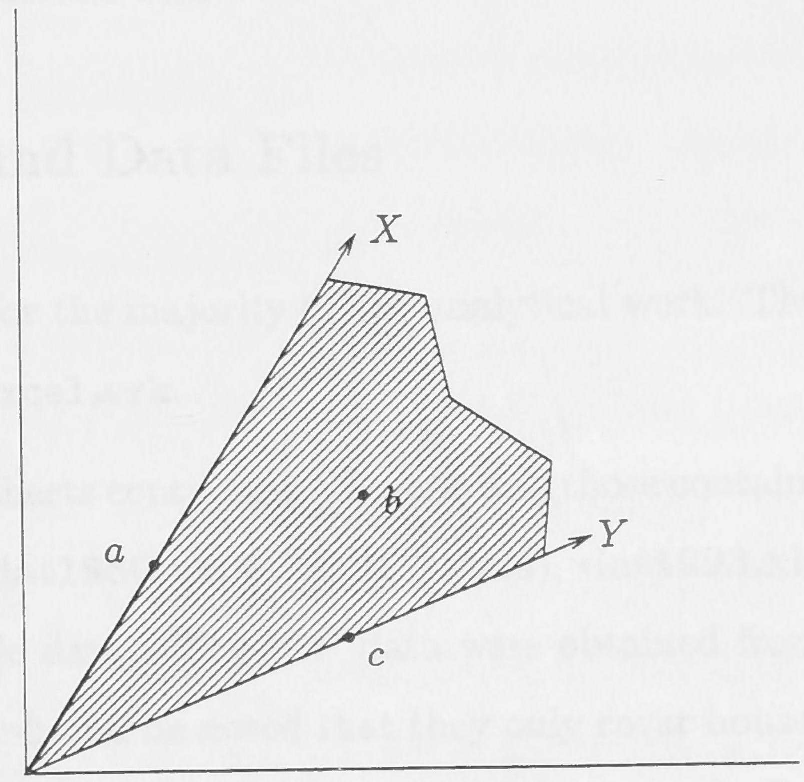

Figure B.6: A convex polyhedral

Appendix C

Programs and Data Files

The empirical analysis in this thesis was conducted using programs written in Visual Basic (VB), C and Gauss. These programs, and the associated data files are described in this Appendix and are available on request from the author. Note that while the programs were written as carefully and clearly as possible, they are not completely "user friendly" and therefore they are released to researchers "as is". In particular, to use these programs one will need to edit the code to suit the task at hand.

C.1 Visual Basic Programs and Data Files

Visual Basic (in Microsoft Excel) was used for the majority of the analytical work. The following is a description of the directory `\excel_wrk`.

There are two types of spreadsheets - spreadsheets containing VB code and those containing data. The data spreadsheets are titled **dat1980.xls** (1980 ICP data), **dat1993.xls** (1993 ICP data) and **examdat.xls** (example data). The ICP data were obtained from the World Bank's *STARS* software, and it should be noted that they only cover household consumption expenditure (not including net expenditure of residents abroad). For further details of the ICP data, see Appendix A.

\harpgarp - subdirectory containing general programs

Harpgarp.xls contains the VB code for data set manipulation, testing of GARP, HARP and AHARP. The procedures within this spreadsheet are designed to work on data contained in the data spreadsheets.

The procedure **Data_mod.Data** creates a subset of price and quantity data and then calculates certain matrices (such as the matrix of log Laspeyres indexes) which are used as input into the revealed preference (RP) tests. These matrices are printed in worksheets in the data spreadsheets. The worksheet *info_t* in the data spreadsheets is used to select the sample of countries for the analysis (column C) and the number of goods and services (column G).

The procedure **Data_mod.Data_affine** performs the same calculations as **Data_mod.Data**, except it works on marginal consumption data - quantities in excess of the minimum subsistence bundle which is contained in the worksheets *gamma* in the data spreadsheets (in sub-directories `\affine` and `\leisure`).

The procedure **GARP_mod.GARP** uses Warshall's algorithm to test for GARP (as per Varian (1982)). Results of the test are printed in the worksheet *GARPlog* in the data spreadsheets.

The procedure **HARP_mod.HARP** uses Warshall's algorithm to test for HARP (as per Varian (1983)). Results of the test are printed in the worksheet *HARPlog* in the data spreadsheets, and the minimum path matrix (and Ideal Afriat Index) are printed in the worksheet *M*. If the procedure **Data_affine** is run before running **HARP_mod.HARP**, then the test will be of AHARP and the worksheet *M* will contain the minimum path matrix defined over marginal quantities and the Ideal Afriat Marginal Index.

The procedure **HARP_searcher_mod.HARP_searcher** is used to iterate through the countries and select the largest set of countries sharing common homothetic preferences. This routine takes a long time in VB, so it has also been implemented in C (see below).

The procedure **Bounds_as_add_mod.Bounds_as_add** prints in the worksheet *bounds_add_log* the true bounds (derived from **M**) as countries are iteratively added to the set satisfying HARP. This was used in Chapter 3.

The procedure **IAMI_bounds_mod.IAMI_bounds** calculates the Ideal Afriat Marginal Index for different γs (found using a C program to be described below) and calculates overall bounds to the this index and the implied rankings. The results are printed in the worksheet *IAMI_bounds*. This was used in Chapter 5.

\Bbc - subdirectory for improved GARP analysis in Chapter 2

The spreadsheet **rp_bbc.xls** contains two modules for conducting improved GARP tests of the representative consumer hypothesis and constructing improved GARP bounds to the bilateral true welfare indexes. The modules were used in the analysis in Chapter 2 and are based on results in Blundell, Browning, and Crawford (1998). The module **cb_H_mod** uses homothetic expansion paths (this replicates the minimum path matrix, and is used only as a test of the BBC algorithms). The module **cb_QHexact_mod** uses quasi-homothetic preferences.

\ChavCox - subdirectory for Afriat envelope analysis in Chapter 2

The spreadsheet **chav_env4.xls** contains three modules for constructing Afriat envelope functions and the implied bounds to the bilateral true welfare indexes. The modules were used in the analysis in Chapter 2 and are based on results in Chavas and Cox (1997). The module **Afr_Uncond_mod** constructs unconditional Afriat inner and outer bound envelope functions when preferences are general. The unconditional Afriat numbers are used as input into the module **Afr_Cond_mod** which constructs inner-bound and outer-bound representations of the expenditure function, and the bounds to the bilateral true indexes implied by these representations. The module **Afr_Uncond_mod_h** constructs the unconditional Afriat numbers under the assumption of homotheticity. This module replicates the minimum path matrix.

\DiewPark - subdirectory for mathematical programs used in Chapter 2

The spreadsheet **Diew85_lp.xls** contains two modules which conduct mathematical program tests of common preferences using results from Diewert and Parkan (1985). The module **GARP_LP_mod** conducts the mathematical program test of common general preferences. The module **HARP_LP_mod** conducts the mathematical program test of common homothetic preferences.

\Var82 - subdirectory for GARP analysis in Chapter 2

There are two spreadsheets containing VB code in this subdirectory.

The spreadsheet **Var82_Afr.xls** contains two modules. The module **Afrcomb_mod** computes conditional Afriat numbers using the combinatorial algorithm (Algorithm 3) in Varian (1982). The module **Afrcomb_h_mod** computes conditional homothetic Afriat numbers using a modified version of the combinatorial algorithm in Varian (1982).

The spreadsheet **Var82_bounds.xls** contains one module. The module **GARP_bounds_mod** calculates the classical and fixed-weight bounds to the bilateral true welfare index and the GARP bounds suggested by Varian (1982).

`\e_effic` - subdirectory for analysis in Chapter 4

There are two spreadsheets containing VB code in this subdirectory.

The spreadsheet **e_effic.xls** contains two modules. The module **e_effic_mod** is used to test HARP with different levels of consumer optimisation error and constructs the minimum path matrix. Note: the procedure **Data_mod.Data** must be run before running **e_effic_mod**. The module **e_effic_searcher_mod** is used to iterate over non-homothetic countries and finds the level of the Afriat efficiency parameter (e^*) necessary so that each country is found to share homothetic preferences. Note: the procedure **Data_mod.Data** must be run before running **e_effic_searcher_mod**.

The spreadsheet **envel.xls** contains two modules. The module **envel_mod_h** is used to construct inner and outer Afriat envelope functions (as per Chavas and Cox (1997)), conditional on the Afriat numbers contained in the minimum path matrix (thus it in fact constructs unconditional inner and outer envelope functions). Note: the procedures **Data_mod.Data** and **HARP_mod.HARP** must be run before running **envel_mod_h**. The module **impute_mod_h** is used to impute utility bounds (using the Afriat envelope functions) for those countries which do not share common homothetic preferences. Note: the procedures **Data_mod.Data** and **HARP_mod.HARP** must be run before running **impute_mod_h**.

`\affine` - subdirectory for analysis in Chapter 5

This subdirectory contains only data spreadsheets. The code for finding the set of γs consistent with AHARP (and the mean γ) was written in C (see below). Once the mean γ has been estimated in C, it is then placed in the worksheet *gamma* in the data spreadsheets and the Ideal Afriat Marginal Index is calculated using the procedures

Data_mod.Data_affine and **HARP_mod.HARP** described above.

\leisure - subdirectory for analysis in Chapter 6

This subdirectory contains only data spreadsheets. The spreadsheet **w_h_dat93.xls** contains the raw data for hours of work and wages, **pop93.xls** contains the calculations for adjusting the 1980 hours of work data for cross-country differences in demographic composition and **Leidat93.xls** contains the worksheets for the calculation of the leisure-inclusive welfare indexes described in Chapter 6.

C.2 C Programs and Data Files

The C programming language was used to conduct some of the iterative routines which took too long to run in VB. The programming was performed using the Linux operating system with the GCC compiler. The following is a discussion of the files in the directory `\C_wrk`. The analysis of the 1980 and 1993 ICP data is separated into the two subdirectories `\80_wrk` and `\93_wrk`.

The program **icp_prog.c** contains the source code (this program is called **icp_prog_93.c** in `\93_wrk`). This program contains several procedures - to run a particular procedure it is necessary to edit the source to make the procedure active (inactive procedures are commented out) and then re-compile the code to create an executable file. All procedures require three input files: *p_k18.dat* (ICP price data on 18 goods), *p_k18.dat* (ICP quantity data on 18 goods) and *subset.dat* (the data file which is used to select which countries are in the analysis). Thus, to select a particular subset of countries for analysis, *subset.dat* must be edited accordingly, and then **icp_prog.c** compiled (with the appropriate procedure activated). It should also be noted that the user has to manually set the variables k and K in **icp_prog.c**. Also note that the Afriat efficiency parameter e^* (labeled e in the program) must be set at a particular value. Note that *p_k19.dat* contains ICP price data on 18 goods, plus wages and *q_k19_l0.dat* contains ICP quantity data on 18 goods, plus leisure (where $T = 24$) and *q_k19_l0_adj.dat* contains ICP quantity data on 18 goods, plus leisure (where $T = 24$ and data have been adjusted for cross-country differences in demographic composition).[1]

[1]Note that the 1993 data includes one extra good (public medical care).

The main procedures in **icp_prog.c** are now described.

HARP_searcher Starting with an initial (arbitrarily selected) subset of four countries, this procedure iterates over the remaining countries and find the maximum number of countries satisfying HARP. The set of countries shown to share common homothetic preferences is printed in the output file *HARPsub.dat.*

AHARP_searcher Starting with an initial (arbitrarily selected) subset of four countries, this procedure iterates over the remaining countries and find the maximum number of countries satisfying AHARP. A particular country is added to the set of countries satisfying AHARP if AHARP is satisfied for at least 10 γs out of a possible 1000 draws from the uniform distribution of potential subsistence bundles. The set of countries satisfying AHARP is printed in the output file *AHARPsub.dat.* Details of the number of random draws of γ performed for each country, and the values of γ which satisfy AHARP are printed in the file *AHARPsearcher.out.* The number of successful draws out of 1000 required for a country to be added to the AHARP set of countries can be changed by altering the parameter gam_n_t_search (initially set to 10).

AHARP_test For the countries selected using the procedure **AHARP_searcher**, there exist an infinite number of subsistence bundles which result in these countries sharing common affine homothetic preferences. The procedure **AHARP_test** is used to collect an "adequate sample" of these consistent γs and then calculate the mean of this set (which is then used in the construction of the Ideal Afriat Marginal welfare index using the VB procedures **Data_affine** and **HARP_mod.HARP**). The number of consistent γs required for an "adequate sample" is arbitrary - it is set to 500 at present, but can be changed by changing the parameter gam_n_t_test. The file *AHARPtest.out* contains the printout the results of **AHARP_test**, including the number of random draws required to collect the sample of consistent γs, the values of the consistent γs, and relevant descriptive statistics such as minimum, maximum, mean.

CalcElasticities The procedure **CalcElasticities** is used to calculate estimated income elasticities of demand. The file *gamma.dat* contains the vector of mean γs (obtained from the procedure **AHARP_test**) which is used as input into **CalcElasticities**, and the file *elast.out* contains the printout of the estimated elasticities.

C.3 Gauss Programs

Gauss was used for calculating the Geary index using the matrix algebra approach suggested by Neary (2000). The file *Geary.prg* in the subdirectory `\gauss_wrk` contains the Gauss code and the example data from Appendix A.

List of Notation

Concepts which are used frequently throughout this thesis have a common notation; hence, in this list, the notation is given only for the first chapter in which it occurs.

Chapter 1	
N	number of countries in the data set
K	number of commodities
$\mathbf{p}^i$	price vector for country i, with typical element p_l^i, $l = 1, \ldots, K$
$\mathbf{q}^i$	quantity vector for country i, with typical element q_l^i, $l = 1, \ldots, K$
$\mathbf{p}.\mathbf{q}$	the inner product of $\mathbf{p}$ and $\mathbf{q}$
Q_{ij}	index expressing welfare of country i relative to that of country j
x	outlay, budget or total expenditure ($= \mathbf{p}.\mathbf{q}$)
x_l	expenditure on good l ($= p_l q_l$)
$\mathbf{q}(x, \mathbf{p})$	Marshallian demand function
$u(\mathbf{q})$	(unobservable) utility function
U	(unobservable) utility level ($= u(\mathbf{q})$)
$\mathbf{h}(U, \mathbf{p})$	Hicksian demand function
$\psi(x, \mathbf{p})$	indirect utility function
$e(U, \mathbf{p})$	expenditure function
$d(U^r, \mathbf{q})$	distance function ($= \max_\delta\{\delta : u(\mathbf{q}/\delta) \geq U^r\}$)
ω_l	budget share for commodity l ($= p_l q_l / x$)
$e_C(U^i, \mathbf{p}^j)$	classical approximation to expenditure function
	($= \min_\mathbf{q} \mathbf{p}^j.\mathbf{q}$ subject to $\mathbf{p}^i.\mathbf{q} = x^i$)
$\boldsymbol{\pi}$	Geary "world prices"
$\boldsymbol{\epsilon}$	Geary "true" exchange rates
$\boldsymbol{\Pi}$	GAIA "world prices"

ϵ	GAIA "true" exchange rates
q_l^{*i}	GAIA imputed quantity

Chapter 2

R^D	directly revealed preferred relation: $\mathbf{q}^i R^D \mathbf{q}$ if $\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}$
P^D	strictly directly revealed preferred relation: $\mathbf{q}^i P^D \mathbf{q}$ if $\mathbf{p}^i.\mathbf{q}^i > \mathbf{p}^i.\mathbf{q}$
R	revealed preferred relation: the transitive closure of R^D
P	strictly revealed preferred relation: $\mathbf{q}^i P \mathbf{q}$ if $\exists$ $\mathbf{q}^j, \mathbf{q}^l$ such that $\mathbf{q}^i R \mathbf{q}^j, \mathbf{q}^j P^D \mathbf{q}^l, \mathbf{q}^l R \mathbf{q}$
$w(\mathbf{q})$	utility function that rationalises the demand data
$RW(\mathbf{q}^0)$	set of observations revealed worse to $\mathbf{q}^0$, $\{\mathbf{q} : \mathbf{q}^0 P \mathbf{q}\}$
$RP(\mathbf{q}^{'})$	set of observations revealed preferred to $\mathbf{q}^{'}$, $\{\mathbf{q} : \mathbf{q} P \mathbf{q}^{'}\}$
$NRW(\mathbf{q}^0)$	set of observations not revealed worse than $\mathbf{q}^0$, $\{\mathbf{q} : \mathbf{p}^i.\mathbf{q} > \mathbf{p}^i.\mathbf{q}^i$ for all $\mathbf{q}^i$ such that $\mathbf{q}^0 R \mathbf{q}^i\}$
W	utility level $(= w(\mathbf{q}))$, or Afriat number
λ	marginal utility of income, or Afriat number
$x \in X$	x belongs to X (x is a member of X)
$x \notin X$	x does not belong to X (x is not a member of X)
$X \setminus Y$	$\{x : x \in X, x \notin Y\}$
$X \cup Y$	$\{x : x \in X, x \in Y\}$
$X \subset Y$	X is contained in Y (X is a subset of Y)
$X \supset Y$	X contains in Y (X is a supset of Y)
$\subseteq$	X is contained in Y or is equivalent to Y
$\mathbf{G}$	matrix summarising relation R^D: $G_{ij} = 1$ if $\mathbf{q}^i R^D \mathbf{q}^j$, 0 otherwise
$\mathbf{H}$	matrix summarising relation R: $H_{ij} = 1$ if $\mathbf{q}^i R \mathbf{q}^j$, 0 otherwise
$\mathbf{P}$	matrix summarising relation P^D: $P_{ij} = 1$ if $\mathbf{q}^i P^D \mathbf{q}^j$, 0 otherwise
$\mathbf{v}^i$	normalised prices $(= \mathbf{p}^i/\mathbf{p}^i.\mathbf{q}^i)$
$e_{max}(U^i, \mathbf{p}^j)$	improved (via RP) upper bound to expenditure function
$e_{min}(U^i, \mathbf{p}^j)$	improved (via RP) lower bound to expenditure function
$P(\mathbf{q}^0)$	$\{\mathbf{q} : u(\mathbf{q}) > u(\mathbf{q}^0)\}$
$m(\mathbf{q}^0, \mathbf{p})$	money metric utility function proposed by Varian (1982) $(= \inf \mathbf{p}.\mathbf{q}$ subject to $\mathbf{q}$ in $P(\mathbf{q}^0))$

$m^+(\mathbf{q}^0, \mathbf{p})$	upper bound on money metric function, (= $\inf \mathbf{p}.\mathbf{q}$ subject to $\mathbf{q}$ in $RP(\mathbf{q}^0)$)
$m^-(\mathbf{q}^0, \mathbf{p})$	lower bound on money metric function, (= $\inf \mathbf{p}.\mathbf{q}$ subject to $\mathbf{q}$ in $NRW(\mathbf{q}^0)$)
$am^+(\mathbf{q}^0, \mathbf{p})$	approximation to $m^+(\mathbf{q}^0, \mathbf{p})$, (= $\min \mathbf{p}.\mathbf{q}^i$ such that $\mathbf{q}^i R \mathbf{q}^0$)
$m^-(\mathbf{q}^0, \mathbf{p})$	approximation to $m^-(\mathbf{q}^0, \mathbf{p})$, (= $\min \mathbf{p}.\mathbf{q}^j$ such that $\mathbf{p}^i.\mathbf{q}^j > \mathbf{p}^i.\mathbf{q}^i$ for all $\mathbf{q}^i$ such that $\mathbf{q}^0 R \mathbf{q}^i$)
$m^{I+}(\mathbf{q}^0, \mathbf{p})$	upper bound on money metric function, calculated using expansion path information
$m^{I-}(\mathbf{q}^0, \mathbf{p})$	lower bound on money metric function, calculated using expansion path information
$CM(\mathbf{q}^0)$	interior of convex hull of $\{\mathbf{q} : \mathbf{q} \geq \mathbf{q}^i, \mathbf{q}^i R \mathbf{q}^0\}$
$\overline{CM}(\mathbf{q}^0)$	closure of $CM(\mathbf{q}^0)$
$EP(\mathbf{p}^i)$	expansion path for country i
$w_I(\mathbf{q}, \boldsymbol{W})$	conditional inner envelope function
$w_O(\mathbf{q}, \boldsymbol{W}, \boldsymbol{\lambda})$	conditional outer envelope function
$w_{UI}(\mathbf{q})$	unconditional inner envelope function
$w_{UO}(\mathbf{q})$	unconditional outer envelope function
$e_I(W^s, \mathbf{p})$	unconditional inner-bound representation of expenditure function
$e_O(W^s, \mathbf{p})$	unconditional outer-bound representation of expenditure function

Chapter 3

Λ	unobservable marginal utility of income
$a(\mathbf{q})$	homothetic utility function that rationalises the demand data
A	homothetic utility level (= $a(\mathbf{q})$), or Afriat number
a	= $\log A$

Chapter 4

$\mathcal{H}$	set of countries which share common homothetic preferences
$\mathcal{R}$	set of countries which do not share common homothetic preferences with countries in $\mathcal{H}$

$a_I(\mathbf{q}, \boldsymbol{A})$	conditional inner homothetic envelope function
$a_O(\mathbf{q}, \boldsymbol{A})$	conditional outer homothetic envelope function
$a_{UI}(\mathbf{q})$	unconditional inner homothetic envelope function
$a_{UO}(\mathbf{q})$	unconditional outer homothetic envelope function
e^*	Afriat efficiency index
$R^D(e^*)$	directly revealed preferred relation (efficiency level e^*): $\mathbf{q}^i R^D(e^*)\mathbf{q}$ if $e^*\mathbf{p}^i.\mathbf{q}^i \geq \mathbf{p}^i.\mathbf{q}$
$R(e^*)$	transitive closure of $R^D(e^*)$

Chapter 5

γ	minimum subsistence consumption bundle
$\mathcal{G}$	set of minimum subsistence consumption bundles
G	number of bundles sampled from $\mathcal{G}$

Chapter 6

l	(supernumerary) leisure time consumed per capita per year
γ_{K+1}	minimum subsistence leisure per capita per year
$u(\mathbf{q}, l)$	leisure-augmented utility function
h	hours of work per capita per year
w	wage rate
m	exogenous non-labour income
T	endowment of time split between labour and leisure
H	total number of hours in year
y	full income, $= \mathbf{p}.\mathbf{q} + wl = wT + m$
$f(\mathbf{p}, w, U^*)$	full income expenditure function, $= \min\limits_{(\mathbf{q},l)}\{\mathbf{p}.\mathbf{q} + wl : u(\mathbf{q}, l) \geq U^*\}$
$\tilde{\mathbf{p}}$	leisure-augmented price bundle ($=\{p_1, \ldots, p_K, w\}$)
$\tilde{\mathbf{q}}$	leisure-augmented consumption vector ($=\{q_1, \ldots, q_K, (H-h)\}$))
$\tilde{\gamma}$	leisure-augmented minimum subsistence consumption bundle
$\tilde{\mathcal{G}}$	set of leisure-augmented minimum subsistence consumption bundles

List of Index Numbers

The following is a summary of the main index numbers which are studied in this thesis. The index numbers are grouped by the first chapter in which they occur.

Chapter 1

$Q_{ij}^{L} = \mathbf{p}^j.\mathbf{q}^i/\mathbf{p}^j.\mathbf{q}^j$	Laspeyres quantity index
$Q_{ij}^{P} = \mathbf{p}^i.\mathbf{q}^i/\mathbf{p}^i.\mathbf{q}^j$	Paasche quantity index
$Q_{ij}^{F} = \sqrt{Q_{ij}^{L}Q_{ij}^{P}}$	Fisher index
$Q_{ij}^{EKS} = \prod_{k=1}^{n}\left(\frac{Q_{ik}^F}{Q_{jk}^F}\right)^{(1/n)}$	EKS index
$Q_{ij}^{T} = \prod_{l=1}^{m}\left(\frac{q_l^i}{q_l^j}\right)^{\frac{\omega_l^i+\omega_l^j}{2}}$	Törnqvist index
$Q_{ij}^{CCD} = \prod_{k=1}^{n}\left(\frac{Q_{ik}^T}{Q_{jk}^T}\right)^{(1/n)}$	CCD index
$Q_{ij}^{A,r} = e(U^i,\mathbf{p}^r)/e(U^j,\mathbf{p}^r)$	Allen Quantity index (at reference prices $\mathbf{p}^r$)
$Q_{ij}^{LA} = e(U^i,\mathbf{p}^j)/e(U^j,\mathbf{p}^j)$	Laspeyres-Allen Quantity index
$Q_{ij}^{PA} = e(U^i,\mathbf{p}^i)/e(U^j,\mathbf{p}^i)$	Paasche-Allen Quantity index
$P_{ij}^{K,r} = e(U^r,\mathbf{p}^i)/e(U^r,\mathbf{p}^j)$	Konüs cost-of-living index (at reference utility U^r)
$P_{ij}^{LK} = e(U^j,\mathbf{p}^i)/e(U^j,\mathbf{p}^j)$	Laspeyres-Konüs true cost-of-living index
$P_{ij}^{PK} = e(U^i,\mathbf{p}^i)/e(U^i,\mathbf{p}^j)$	Paasche-Konüs true cost-of-living index
$Q_{ij}^{K,r} = \frac{x^i}{x^j P_{ij}^{K,r}}$	Konüs Quantity index (at U^r)
$Q_{ij}^{MQ,r} = \frac{d(U^r,\mathbf{q}^i)}{d(U^r,\mathbf{q}^j)}$	Malmquist Quantity index (at U^r)
$Q_{ij}^{LMQ} = \frac{d(U^j,\mathbf{q}^i)}{d(U^j,\mathbf{q}^j)}$	Laspeyres-Malmquist Quantity index
$Q_{ij}^{PMQ} = \frac{d(U^i,\mathbf{q}^i)}{d(U^i,\mathbf{q}^j)}$	Paasche-Malmquist Quantity index
$Q_{ij}^{LM} = \frac{d(U^j,\mathbf{q}^j)}{d(U^i,\mathbf{q}^j)}$	Laspeyres-Malmquist Welfare index
$Q_{ij}^{PM} = \frac{d(U^j,\mathbf{q}^i)}{d(U^i,\mathbf{q}^i)}$	Paasche-Malmquist Welfare index
$Q_{ij}^{CL} = e_C(U^i,\mathbf{p}^j)/x^j$	classical lower bound to Q_{ij}^{LA}

$Q_{ij}^{CU} = x^i/e_C(U^j, \mathbf{p}^i)$ classical upper bound to Q_{ij}^{PA}

B base-weighted true index

C current-weighted true index

A unique true index

Chapter 2

$\mathbf{L}$ matrix of log Q_{ij}^L: $\{L_{ij}\} = \log(\mathbf{p}^i.\mathbf{q}^j/\mathbf{p}^i.\mathbf{q}^i)$

$\mathbf{C}$ matrix of log classical upper bounds to Q_{ij}^{PA}:
$\{C_{ij}\} = \log[x^j/e_C(U^i, \mathbf{p}^j)]$

$\mathbf{B}^{RP}$ matrix of log RP upper bounds to Q_{ij}^{LA}: $\{B_{ij}^{RP}\} = \log[e_{max}(U^j, \mathbf{p}^i)/x^i]$

$\mathbf{C}^{RP}$ matrix of log RP upper bounds to Q_{ij}^{PA}: $\{C_{ij}^{RP}\} = \log[x^j/e_{min}(U^i, \mathbf{p}^j)]$

$\mathbf{B}^G$ matrix of log GARP upper bounds to Q_{ij}^{LA}: $\{B_{ij}^G\} = \log[m^+(\mathbf{q}^j, \mathbf{p}^i)/x^i]$

$\mathbf{C}^G$ matrix of log GARP upper bounds to Q_{ij}^{PA}: $\{C_{ij}^G\} = \log[x^j/m^-(\mathbf{q}^i, \mathbf{p}^j)]$

$\mathbf{B}^{IG}$ matrix of log improved GARP upper bounds to Q_{ij}^{LA}:
$\{B_{ij}^{IG}\} = \log[m^{I+}(\mathbf{q}^j, \mathbf{p}^i)/x^i]$

$\mathbf{C}^{IG}$ matrix of log improved GARP upper bounds to Q_{ij}^{PA}:
$\{C_{ij}^{IG}\} = \log[x^j/m^{I-}(\mathbf{q}^i, \mathbf{p}^j)]$

$\mathbf{B}^A$ matrix of log Afriat envelope upper bounds to Q_{ij}^{LA}:
$\{B_{ij}^A\} = \log[e_I(W^j, \mathbf{p}^i)/x^i]$

$\mathbf{C}^A$ matrix of log Afriat envelope upper bounds to Q_{ij}^{PA}:
$\{C_{ij}^A\} = \log[x^j/e_O(W^i, \mathbf{p}^j)]$

Chapter 3

$\mathbf{M}$ minimum path matrix, $= \min_{k,\ldots,m}\{L_{ij}, (L_{ik} + L_{kl} + \ldots + L_{mj})\}$

$\mathbf{a}^+$ vector of column means of $\mathbf{M}$

$\mathbf{a}^-$ vector of negative row means of $\mathbf{M}$

$\mathbf{a}^*$ Ideal Afriat Index ($\equiv (\mathbf{a}^+ + \mathbf{a}^-)/2$)

Chapter 4

$\mathbf{M}^a$ augmented minimum path matrix

$\mathbf{a}^{a+}$ vector of column means of $\mathbf{M}^a$

$\mathbf{a}^{a-}$ vector of negative row means of $\mathbf{M}^a$

$\mathbf{a}^{a*}$ Augmented Ideal Afriat Index ($\equiv (\mathbf{a}^{a+} + \mathbf{a}^{a-})/2$)

$\mathbf{L}^e$ matrix of log Q^L_{ij} incorporating Afriat efficiency level e: $\{L^e_{ij}\} = \log(\mathbf{v}^i.\mathbf{q}^j/e)$

$\mathbf{M}^e$ minimum path matrix incorporating Afriat efficiency level e

$\mathbf{a}^+(e)$ vector of column means of $\mathbf{M}^e$

$\mathbf{a}^-(e)$ vector of negative row means of $\mathbf{M}^e$

$\mathbf{a}^*(e)$ Approximate Ideal Afriat Index ($\equiv (\mathbf{a}^+(e) + \mathbf{a}^-(e))/2$)

Chapter 5

$Q^{AM,r}_{ij} = \frac{e(U^i,\mathbf{p}^r)-a(\mathbf{p}^r)}{e(U^j,\mathbf{p}^r)-a(\mathbf{p}^r)}$ Allen Marginal Quantity index (at reference prices $\mathbf{p}^r$)

$Q^{LAM}_{ij} = \frac{e(U^i,\mathbf{p}^j)-a(\mathbf{p}^j)}{e(U^j,\mathbf{p}^j)-a(\mathbf{p}^j)}$ Laspeyres-Allen Marginal Quantity index

$Q^{PAM}_{ij} = \frac{e(U^i,\mathbf{p}^i)-a(\mathbf{p}^i)}{e(U^j,\mathbf{p}^i)-a(\mathbf{p}^i)}$ Paasche-Allen Marginal Quantity index

$Q^{LM}_{ij} = \frac{\mathbf{p}^j.(\mathbf{q}^i-\gamma^j)}{\mathbf{p}^j.(\mathbf{q}^j-\gamma^j)}$ Laspeyres Marginal Quantity index

$Q^{PM}_{ij} = \frac{\mathbf{p}^i.(\mathbf{q}^i-\gamma^i)}{\mathbf{p}^i.(\mathbf{q}^j-\gamma^i)}$ Paasche Marginal Quantity index

$\mathbf{L}^m$ matrix of log Q^{LM}_{ij}: $\{L^m_{ij}\} = \log Q^{LM}_{ji}$

$\mathbf{M}^m$ minimum path matrix defined over marginal quantities

$\mathbf{a}^+_m$ vector of column means of $\mathbf{M}^m$

$\mathbf{a}^-_m$ vector of negative row means of $\mathbf{M}^m$

$\mathbf{a}^*_m$ Ideal Afriat Marginal Index ($\equiv (\mathbf{a}^+_m + \mathbf{a}^-_m)/2$)

Chapter 6

$Y^{A,r}_{ij} = \frac{f(\mathbf{p}^r,w^r,u(\mathbf{q}^i,l^i))}{f(\mathbf{p}^r,w^r,u(\mathbf{q}^j,l^j))}$ Allen real full income index (at reference prices $\mathbf{p}^r$)

$Y^{LA}_{ij} = \frac{f(\mathbf{p}^j,w^j,u(\mathbf{q}^i,l^i))}{y^j}$ Laspeyres-Allen full income index

$Y^{PA}_{ij} = \frac{y^i}{f(\mathbf{p}^i,w^i,u(\mathbf{q}^j,l^j))}$ Paasche-Allen full income index

$Y^{L}_{ij} = \frac{\mathbf{p}^j\mathbf{q}^i+w^il^i}{y^j}$ Laspeyres full income index

$Y^{P}_{ij} = \frac{y^i}{\mathbf{p}^i\mathbf{q}^j+w^il^j}$ Paasche full income index

$Y^{AM,r}_{ij} = \frac{f(U^i,\tilde{\mathbf{p}}^r)-a(\tilde{\mathbf{p}}^r)}{f(U^j,\tilde{\mathbf{p}}^r)-a(\tilde{\mathbf{p}}^r)}$ Allen marginal real full income index (at reference prices $\mathbf{p}^r$)

$Y^{LAM}_{ij} = \frac{f(U^i,\tilde{\mathbf{p}}^j)-a(\tilde{\mathbf{p}}^j)}{f(U^j,\tilde{\mathbf{p}}^j)-a(\tilde{\mathbf{p}}^j)}$ Laspeyres-Allen marginal real full income index

$Y^{PAM}_{ij} = \frac{f(U^i,\tilde{\mathbf{p}}^i)-a(\tilde{\mathbf{p}}^i)}{f(U^j,\tilde{\mathbf{p}}^i)-a(\tilde{\mathbf{p}}^i)}$ Paasche-Allen marginal real full income index

$\tilde{Q}^{LM}_{ij} = \frac{\tilde{\mathbf{p}}^j.(\tilde{\mathbf{q}}^i - \tilde{\gamma}^j)}{\tilde{\mathbf{p}}^j.(\tilde{\mathbf{q}}^j - \tilde{\gamma}^j)}$	Laspeyres leisure-augmented marginal quantity index
$\tilde{Q}^{PM}_{ij} = \frac{\tilde{\mathbf{p}}^i.(\tilde{\mathbf{q}}^i - \tilde{\gamma}^i)}{\tilde{\mathbf{p}}^i.(\tilde{\mathbf{q}}^j - \tilde{\gamma}^i)}$	Paasche leisure-augmented marginal quantity index
$\tilde{\mathbf{L}}^m$	matrix of $\log \tilde{Q}^{LM}_{ij}$: $\{\tilde{L}^m_{ij}\} = \log \tilde{Q}^{LM}_{ji}$
$\widetilde{\mathbf{M}}^m$	minimum path matrix defined over leisure-agumented marginal quantities
$\tilde{\mathbf{a}}^+_m$	vector of column means of $\widetilde{\mathbf{M}}^m$
$\tilde{\mathbf{a}}^-_m$	vector of negative row means of $\widetilde{\mathbf{M}}^m$
$\tilde{\mathbf{a}}^*_m$	Ideal Afriat Marginal Index $(\equiv (\tilde{\mathbf{a}}^+_m + \tilde{\mathbf{a}}^-_m)/2)$

Author Index

Subject Index

Bibliography

ABBOTT, M., AND O. ASHENFELTER (1976): "Labor Supply, Commodity Demand, and the Allocation of Time," *Review of Economic Studies*, 43, 389–412.

AFRIAT, S. (1967): "The Construction of a Utility Function from Expenditure Data," *International Economic Review*, 8(1), 67–77.

——— (1972): "The Theory of International Comparisons of Real Income and Prices," in *International Comparisons of Prices and Output*, ed. by D. Daly. National Bureau of Economic Research, New York.

——— (1976): *The Combinatorial Theory of Demand*. Input-Output Publishing Company, London.

——— (1977): *The Price Index*. Cambridge University Press, Cambridge.

——— (1981): "On the Constructibility of Consistent Price Indices Between Several Periods Simultaneously," in *Essays in the Theory and Measurement of Consumer Behaviour*, ed. by A. Deaton. Cambridge University Press, Cambridge.

——— (1987): *Logic of Choice and Economic Theory*. Clarendon Press, Oxford.

BALK, B. (1996): "A Comparison of Ten Methods for Multilateral International Prices and Volume Comparison," *Journal of Official Statistics*, 12, 199–222.

BANKER, R., AND A. MAINDIRATTA (1988): "Nonparametric Analysis of Technical and Allocative Efficiencies in Production," *Econometrica*, 56, 1315–1332.

BANKS, J., R. BLUNDELL, AND A. LEWBEL (1997): "Quadratic Engel Curves and Consumer Demand," *Review of Economics and Statistics*, 79, 527–539.

BARNETT, W. (1981): *Consumer Demand and Labor Supply*. Elsevier North-Holland, Amsterdam.

BAYE, M., AND D. BLACK (1992): "Income Taxation, Labor Supply and the Theory of Income-based Cost-of-Living Indices," *European Economic Review*, 36, 83–100.

BLACKORBY, C., R. BOYCE, AND R. RUSSELL (1978): "Estimation of Demand Systems Generated by the Gorman Polar Form: A Generalization of the S-Branch Utility Tree," *Econometrica*, 46(2), 345–363.

BLUNDELL, R., M. BROWNING, AND I. CRAWFORD (1998): "Nonparametric Engel Curves and Revealed Preference," University College London, Discussion Paper 98-08.

Byushgens, S. (1925): "Sur une classe des hypersurfaces; à propos de 'l'index idéal' de M. Irving Fischer," *Recueil Mathématique (Moscow)*, 32, 625–631.

Caves, D., L. Christensen, and W. Diewert (1982): "Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers," *Economic Journal*, 92(365), 73–86.

Chavas, J., and T. Cox (1997): "On Nonparametric Demand Analysis," *European Economic Review*, 41, 75–95.

Cleeton, D. (1982): "The Theory of Real Wage Indices," *American Economic Review*, 70, 214–255.

Coles, J., and P. Harte-Chen (1985): "Real Wage Indices," *Journal of Labor Economics*, 3(3), 317–36.

Deaton, A., and J. Muellbauer (1980): *Economics and Consumer Behaviour*. Cambridge University Press, Cambridge.

Diewert, W. (1973): "Afriat and Revealed Preference Theory," *Review of Economic Studies*, 40(3), 419–26.

——— (1976): "Exact and Superlative Index Numbers," *Journal of Econometrics*, 4, 115–145.

——— (1978): "Superlative Index Numbers and Consistency in Aggregation," *Econometrica*, 47, 883–900.

——— (1981): "The Economic Theory of Index Numbers: a Survey," in *Essays in the Theory and Measurement of Consumer Behaviour*, ed. by A. Deaton. Cambridge University Press, Cambridge.

——— (1983): "The Theory of the Cost-of-Living Index and the Measurement of Welfare Change," in *Price Level Measurement*, ed. by W. Diewert, and C. Montmarquette. Statistics Canada, Ottawa, Canada.

——— (1987): "Index Numbers," in *The New Palgrave: A Dictionary of Economics, Vol. 2*, ed. by J. Eatwell, M. Milgate, and P. Newman. Macmillan, London.

——— (1992): "Fisher Ideal Output, Input and Productivity Indexes Revisited," *Journal of Productivity Analysis*, 3(3), 211–48.

——— (1996): "Axiomatic and Economic Approaches to International Comparisons," NBER Working Paper No. 5559.

Diewert, W., and C. Parkan (1985): "Tests for the Consistency of Consumer Data," *Journal of Econometrics*, 30, 127–147.

Dixit, A. (1990): *Optimization in Economic Theory*. Oxford University Press, Oxford.

Dowrick, S., Y. Dunlop, and J. Quiggin (1995): "Social Indicators and Comparisons of Living Standards," mimeograph, Australian National University.

Dowrick, S., and J. Quiggin (1993): "Australia, Japan and the OECD: GDP Rankings and Revealed Preference," *Australian Economic Review*, 0(101), 21–33.

——— (1994): "International Comparisons of Living Standards and Tastes: A Revealed Preference Approach," *American Economic Review*, 84, 332–41.

——— (1997): "True Measures of GDP and Convergence," *American Economic Review*, 87(1), 41–64.

DRECHSLER, L. (1973): "Weighting of Index Numbers in Multilateral International Comparisons," *Review of Income and Wealth*, 19, 17–34.

EICHHORN, W., AND J. VOELLER (1976): *Theory of the Price Index: Fisher's Test Approach and Generalizations.* Springer-Verlag, Berlin, Heidelburg and New York.

EISNER, R. (1988): "Extended Accounts for National Income and Product," *Journal of Economic Literature*, 26(4), 1611–84.

ELTETÖ, O., AND P. KÖVES (1964): "On a Problem of Index Number Computation Relating to International Comparisons," *Statisztikai Szemle*, 42, 507–518, [in Hungarian].

FISHER, I. (1922): *The Making of Index Numbers.* Houghton Mifflin, Boston.

FRISCH, R. (1936): "Annual Survey of General Economic Theory: The Problem of Index Numbers," *Econometrica*, 4, 1–36.

GEARY, R. (1958): "A Note on the Comparison of Exchange rates and Purchasing Power Between Countries," *Journal of the Royal Statistical Society (Series A)*, 121, 97–99.

GERSCHENKRON, A. (1951): "A Dollar Index fo Soviet Machinery Output, 1927-28 to 1938," Rand Corporation, Santa Monica, CA.

GINI, C. (1931): "On the Circular Test of Index Numbers," *International Review of Statistics*, 9(2), 3–25.

GREENE, W. (1993): *Econometric Analysis.* Macmillan, New York.

HILL, R. (1997): "A Taxonomy of Multilateral Methods for Making International Comparisons," *Review of Income and Wealth*, 43, 49–69.

——— (2000): "Measuring Substitution Bias in International Comparisons Based on Additive Purchasing Power Parity Methods," *European Economic Review*, 44, 145–162.

HOUTHAKKER, H. (1950): "Revealed Preference and the Utility Function," *Economica*, 17, 159–174.

HOUTMAN, M., AND J. MAKS (1987): "The Existence of Homothetic Utility Functions Generating Dutch Consumer Data," University of Groningen, Groningen.

JEHLE, G. (1991): *Advanced Microeconomic Theory.* Prentice-Hall International, Englewood Cliffs.

KHAMIS, S. (1970): "Properties and Conditions for the Existence of a New Type of Index Numbers," *Sankhyā*, 32, 81–98.

——— (1972): "A New System of Index Numbers for National and International Purposes," *Journal of the Royal Statistical Society (Series A)*, 135, 96–121.

Knoblauch, V. (1992): "A Tight Upper Bound on the Money Metric Utility Function," *The American Economic Review*, 82(3), 660–665.

Kokoski, M. (1987): "Indices of Household Welfare and the True Value of Leisure Time," *The Review of Economics and Statistics*, pp. 83–89.

Konüs, A., and S. Byushgens (1926): "On the Problem of the Purchasing Power of Money," *Voprosi Konyunkturi*, II(1), 151–172.

Kuznets, S. (ed.) (1952): *Income and Wealth of the United States, Trends and Structures*, Income and Wealth Series II, International Association for Research in Income and Wealth. Bowes and Bowes, Cambridge, England.

Lerner, A. (1935): "A Note on the Theory of Price Index Numbers," *Review of Economic Studies*, 3, 50–56.

Lloyd, P. (1979): "Constant-Utility Index Numbers of Real Wages: Comment," *American Economic Review*, 69(4), 682–685.

Lluch, C., A. Powell, and R. Williams (1977): *Patterns in Household Demand and Saving*. Oxford University Press, New York.

Manser, M., and R. McDonald (1988): "An Analysis of Substitution Bias in Measuring Inflation, 1959-85," *Econometrica*, 56(4), 909–930.

Neary, J. (1996): "Theoretical Foundations of the "Geary Method" for International Comparisons of Purchasing Power and Real Incomes," *Economic and Social Review*, 27(2), 161–179.

——— (2000): "True Multilateral Indexes for International Comparisons of Real Income: Theory and Empirics," mimeograph: University College Dublin; Ecole Polytechnique, Paris; and CEPR.

Neary, J., and B. Gleeson (1997): "Comparing the Wealth of Nations: Reference Prices and Multilateral Real Income Indexes," *Economic and Social Review*, 28, 401–421.

Nuxoll, D. (1994): "Differences in Relative Prices and International Differences in Growth Rates," *American Economic Review*, 84, 1423–1436.

Patterson, K. (1991): "A Non-parametric Analysis of Personal Sector Decisions on Consumption, Liquid Assets and Leisure," *Economic Journal*, 101, 1103–1116.

Pencavel, J. (1977): "Constant-Utility Index Numbers of Real Wages," *American Economic Review*, 67, 91–100.

——— (1979): "Constant-Utility Index Numbers of Real Wages: Revised Estimates," *American Economic Review*, 69, 240–43.

Pollak, R. (1971): "The Theory of the Cost of Living Index," Research Discussion Paper No. 11, Bureau of Labor Statistics, Washington, D.C.

Richter, M. (1966): "Revealed Preference Theory," *Econometrica*, 34, 635–645.

Riddell, W. (1983): "Leisure Time and the Measurement of Economic Welfare," in *Price Level Measurement: Proceedings from a conference sponsored by Statistics Canada*, ed. by W. Diewert, and C. Montmarquette.

Samuelson, P. (1948): "Consumption Theory in Terms of Revealed Preference," *Economica*, 15, 243–253.

——— (1956): "Social Indifference Curves," *Quarterly Journal of Economics*, 70(1), 1–22.

——— (1964): "Principles of Efficiency - Comment," *American Economic Review*, 54(3), 93–96.

——— (1974): "Complementarity: An Essay on the 40th Anniversary of the Hicks-Allen revolution in Demand Theory," *Journal of Economic Literature*, 64(4), 1255–1289.

Samuelson, P., and S. Swamy (1974): "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis," *The American Economic Review*, 64(4), 566–593.

Selvanathan, E., and D. P. Rao (1994): *Index Numbers: A Stochastic Approach.* University of Michigan Press, Ann Arbor.

Swofford, J., and G. Whitney (1987): "Nonparametric Tests of Utility Maximisation and Weak Separability for Consumption, Leisure and Money," *Review of Economics and Statistics*, 69, 458–64.

——— (1994): "A Revealed Preference Test for Weakly Separable Utility Maximization with Incomplete Adjustment," *Journal of Econometrics*, 60, 234–249.

Szulc, B. (1964): "Index Numbers for Multilateral Regional Comparisons," *Przeglad Statystyczny*, 3, 239–254, [in Polish].

Varian, H. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50(4), 945–73.

——— (1983): "Non-Parametric Tests of Consumer Behaviour," *Review of Economic Studies*, 50(1), 99–110.

——— (1984): "Social Indifference Curves and Aggregate Demand," *Quarterly Journal of Economics*, 99(3), 403–414.

——— (1985): "Non-Parametric Analysis of Optimizing Behavior with Measurement Error," *Journal of Econometrics*, 30, 445–458.

——— (1990): "Goodness-of-Fit in Optimizing Models," *Journal of Econometrics*, 46, 125–140.

——— (1993): "Goodness-of-Fit for Revealed Preference Tests," mimeograph.

Wald, A. (1939): "A New Formula for the Index of Cost of Living," *Econometrica*, 7, 319–331.